PSY9511: Seminar 1

Introduction to machine learning

Esten H. Leonardsen

07.11.24

UNIVERSITY
OF OSLO

Outline

Plan for the day

- Round of introductions
- Course information
- Introduction to machine learning
- Presentation of assignment 1

Teacher

Esten Høyland Leonardsen

- Master's degree in Informatics: Programming and Networks
- PhD in Psychology, deep learning applied to neuroimaging data
- Experience as a data scientist and programmer from the industry and various start-ups
- Post-doc at the center for Cognitive psychology, Neuroscience and Neuropsychology
- Chief Scientific Officer at baba.vision
- Interests: Deep learning, explainable artificial intelligence, mental health, neuroimaging

Students

What I want to know about you

- What's your name?
- What department/section are you from?
- What's your research project about?
- Do you have experience with machine learning and/or programming?
- What do you hope to learn from this course? (e.g. specific applications in your research, a theoretical understanding of machine learning, following and contributing to the public discourse, a future job in data science, …)

About the course

Canvas

- All relevant announcements will be made on Canvas (e.g. changes to assignments, lectures, interesting reading material etc.)
- Lecture slides and notebooks from live coding will be put on Canvas before/after a lecture

About the course

Curriculum

- The course relies on the book ”An Introduction to Statistical Learning”, available at https://www.statlearning.com/
    - Only some chapters will be used, they are posted on Canvas under each Lecture module
    - Although we won’t be relying much on the exercises i **highly recommend** looking into them yourselves
- I will add some scientific publications to the curriculum list as we go, depending on your preferences and interests

About the course

Exercises

- The course has no exam, but six mandatory exercises you will need to pass
  - Mostly practical coding, with some reflection
  - Given with a **hard** deadline, unless there is a good reason for an extension
  - Can be delivered multiple times based on feedback (but the first must be in time for the original deadline)
- Exercises 1-4 and 6 are mostly small and related to specific content of the preceding lecture, while 5 is a bit larger
- You should hand in runnable code (e.g. a Jupyter notebook, a python script, an R script, Rmarkdown etc.), not code copied into a Word document or a pdf

About the course

Generative artificial intelligence (e.g. ChatGPT)

- You are allowed to use generative AI in the assignments, but you must state where and how
  - Be critical, you should be able to understand and explain **all** the code you hand in

About the course

Lectures

- Goal is to show you the underlying theory in an intuitive manner
- $\sim$2 hours of lecturing, $\sim$1 hour for individual work/help with assignments
  - You will have to practice what you learn yourself
- Will try to make lectures interactive, and do live coding where possible

About the course

Course plan

1. Introduction to machine learning
2. Basics of regression and classification
3. Variable selection and regularization
4. Model selection, validation, and testing
5. Non linearity: Splines and tree-based methods
6. Unsupervised learning
7. Deep learning and image processing
8. Language processing

Introduction to machine learning

UNIVERSITY
OF OSLO

Introduction

Key terminology:

- Statistical learning: A set of tools (often called models) for understanding data

Key terminology:

- Statistical learning: A set of tools (often called models) for understanding data
- Supervised learning: We know what task we want to solve
- Unsupervised learning: We don't know what task we want to solve (or we don't have the data we need to solve it)

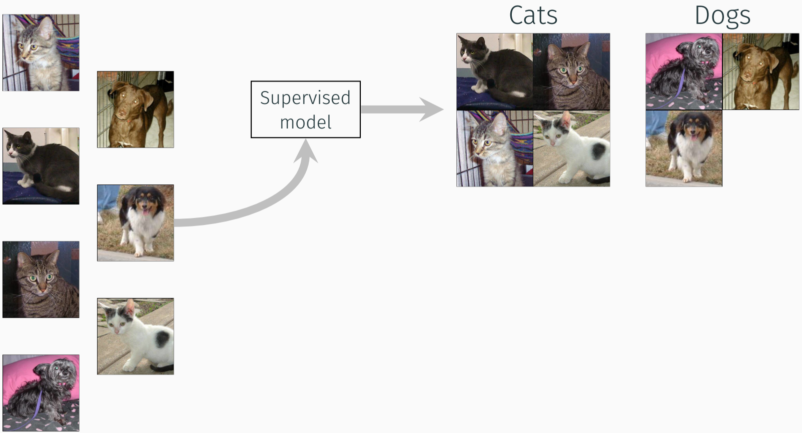
Introduction



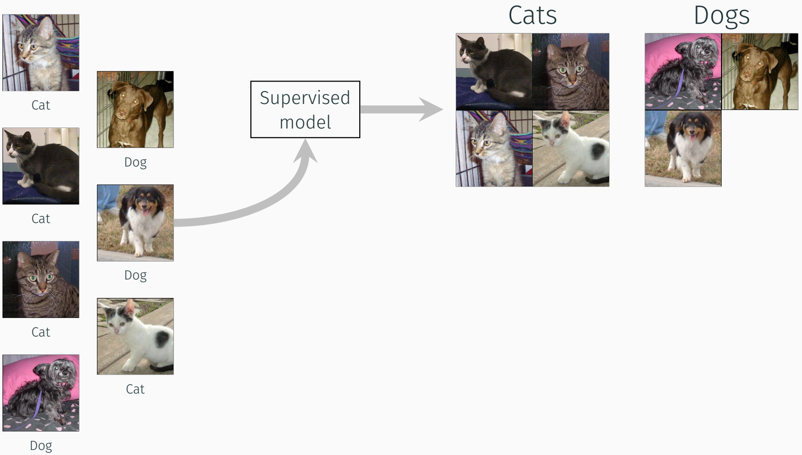


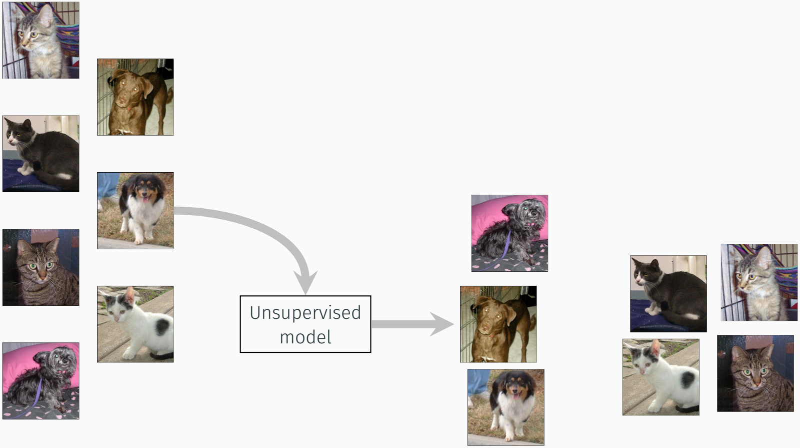


Introduction

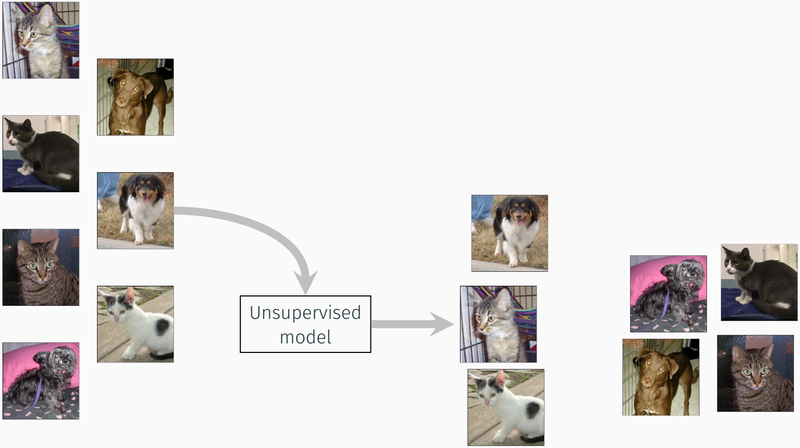


Introduction




Unsupervised
model

Introduction



Introduction: Supervised learning

name	year	cylinders	horsepower	weight	mpg
Chevrolet Chevelle Malibu	1970	8	130	3504	18
Buick Skylark 320	1980	4	165	3693	15
Plymouth Satellite	1971	8	150	3436	18
AMC Rebel SST	1975	4	150	3433	16
Ford Torino	1978	8	140	3449	17

Prerequisites

- A dataset representing a given population

Introduction: Supervised learning

name	year	cylinders	horsepower	weight	mpg
Chevrolet Chevelle Malibu	1970	8	130	3504	18
Buick Skylark 320	1980	4	165	3693	15
Plymouth Satellite	1971	8	150	3436	18
AMC Rebel SST	1975	4	150	3433	16
Ford Torino	1978	8	140	3449	17

Prerequisites

- A dataset representing a given population

Introduction: Supervised learning

name	year	cylinders	horsepower	weight	mpg
Chevrolet Chevelle Malibu	1970	8	130	3504	18
Buick Skylark 320	1980	4	165	3693	15
Plymouth Satellite	1971	8	150	3436	18
AMC Rebel SST	1975	4	150	3433	16
Ford Torino	1978	8	140	3449	17

Prerequisites

- A dataset representing a given population
- A response-variable y that we want to predict

Introduction: Supervised learning

name	year	cylinders	horsepower	weight	mpg
Chevrolet Chevelle Malibu	1970	8	130	3504	18
Buick Skylark 320	1980	4	165	3693	15
Plymouth Satellite	1971	8	150	3436	18
AMC Rebel SST	1975	4	150	3433	16
Ford Torino	1978	8	140	3449	17

Prerequisites

- A dataset representing a given population
- A response-variable y that we want to predict
- A set of predictors X that we can use to predict y

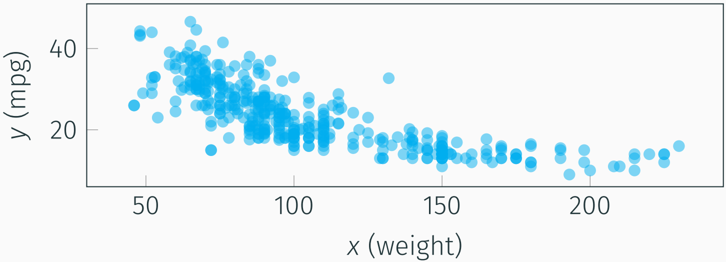
Introduction: Supervised learning

name	year	cylinders	horsepower	weight	mpg
Chevrolet Chevelle Malibu	1970	8	130	3504	18
Buick Skylark 320	1980	4	165	3693	15
Plymouth Satellite	1971	8	150	3436	18
AMC Rebel SST	1975	4	150	3433	16
Ford Torino	1978	8	140	3449	17

Prerequisites

- A dataset representing a given population
- A response-variable y that we want to predict
- A set of predictors X that we can use to predict y
- An **assumed** relationship between X and y that can be described by an unknown function f, such that $y = f(X) + \epsilon$

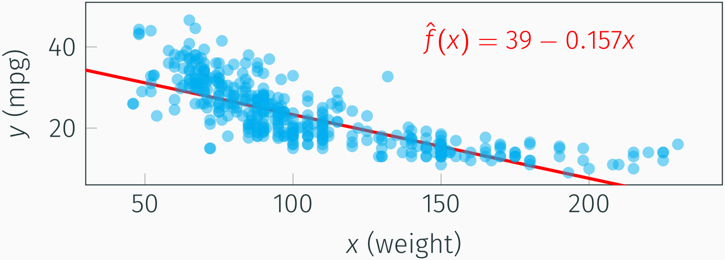
Introduction: Supervised learning



Estimation (or training the model)

- We have assumed that $y = f(X) + \epsilon$, but don't know f

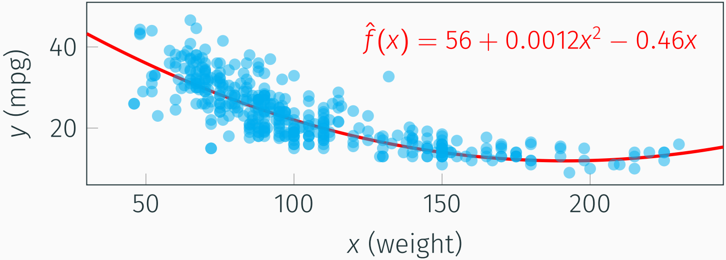
Introduction: Supervised learning



Estimation (or training the model)

- We have assumed that $y = f(X) + \epsilon$, but don't know f
- We produce an estimate $\hat{f}$

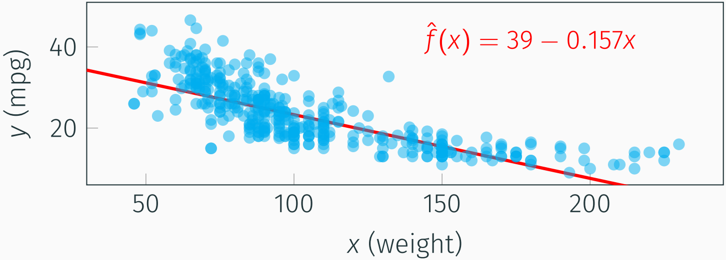
Introduction: Supervised learning



Estimation (or training the model)

- We have assumed that $y = f(X) + \epsilon$, but don't know f
- We produce an estimate $\hat{f}$

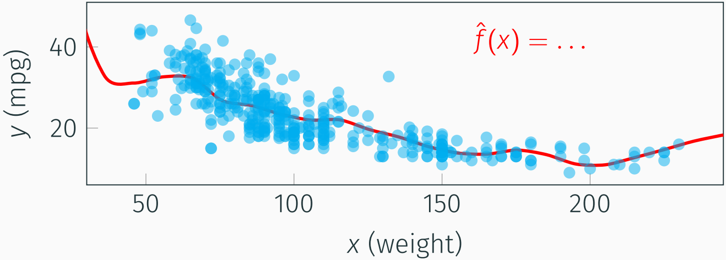
Introduction: Supervised learning



Estimation (or training the model)

- We have assumed that $y = f(X) + \epsilon$, but don't know f
- We produce an estimate $\hat{f}$
- Parametric models: $\hat{f}$ has a simple form
  - $\hat{f}(x) = \beta_0 + \beta_1 x$

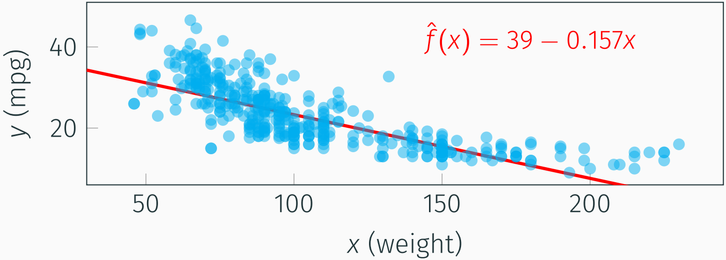
Introduction: Supervised learning



Estimation (or training the model)

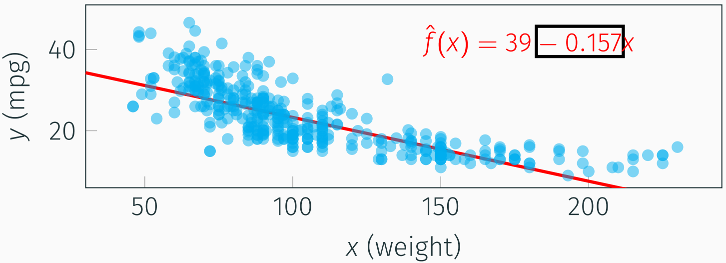
- We have assumed that $y = f(X) + \epsilon$, but don't know f
- We produce an estimate $\hat{f}$
- Parametric models: $\hat{f}$ has a simple form
  - $\hat{f}(x) = \beta_0 + \beta_1 x$
- Non-parametric models: $\hat{f}$ relies directly on the data

Introduction: Supervised learning



Inference: Understanding the relationship between the predictors and the response

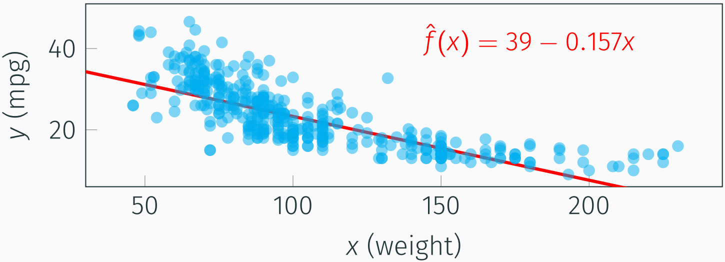
Introduction: Supervised learning



Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?

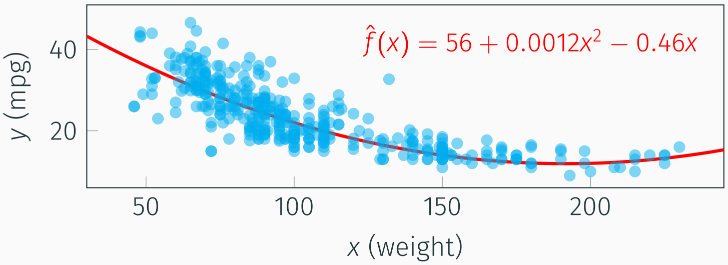
Introduction: Supervised learning



Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?
- What is the functional form?

Introduction: Supervised learning



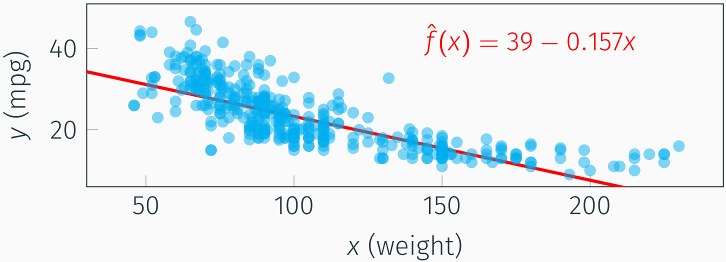
Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?
- What is the functional form?

Prediction: Predicting the response for new observations

- Plugging new values X into $\hat{f}(X)$

Introduction: Supervised learning



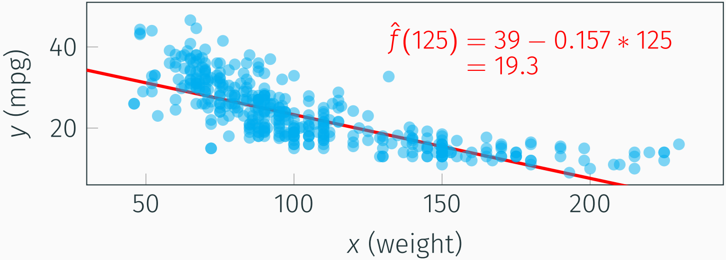
Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?
- What is the functional form?

Prediction: Predicting the response for new observations

- Plugging new values X into $\hat{f}(X)$

Introduction: Supervised learning



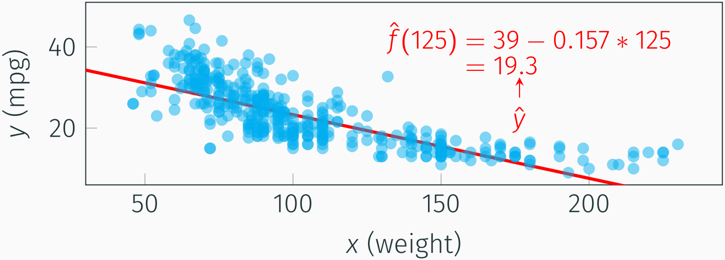
Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?
- What is the functional form?

Prediction: Predicting the response for new observations

- Plugging new values X into $\hat{f}(X)$

Introduction: Supervised learning



Inference: Understanding the relationship between the predictors and the response

- How does individual features relate to the response?
- What is the functional form?

Prediction: Predicting the response for new observations

- Plugging new values X into $\hat{f}(X)$