Applied Machine Learning lecture 1

Introduction to Applied Machine Learning

Pascal Mettes – University of Amsterdam

Lecturer of today: Thomas Mensink - Google

Agenda of the day

Welcome

- Introduction
- Course and syllabus

Regression I

Welcome

PhD Computer Science @ Grenoble (France)

PostDoc researcher @ UvA

Assistant Professor @ UvA

Research Scientist @ Google

Research topics: 3D, Deep Learning, large scale image classification, machine learning, learning without examples.

Lecturer

PhD Computer Science @ UvA

Visiting Researcher @ Columbia University, USA

PostDoc researcher @ UvA

Assistant professor @ UvA

Research topics: action recognition, action localization, deep learning, learning from limited supervision.


퀄리티가 다르다!
대한
극장


ICCV 2019
Seoul, Korea
ICCV 2019
International Conference on Computer Vision 2019

Hi


8
17
88
13
임시


Attendance History (32 years of ICCV)

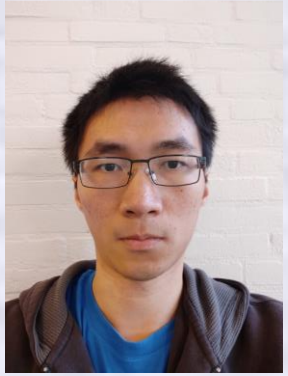
Other people of the course

Cristian Rodriguez Rivero

Ivan Sosnovik

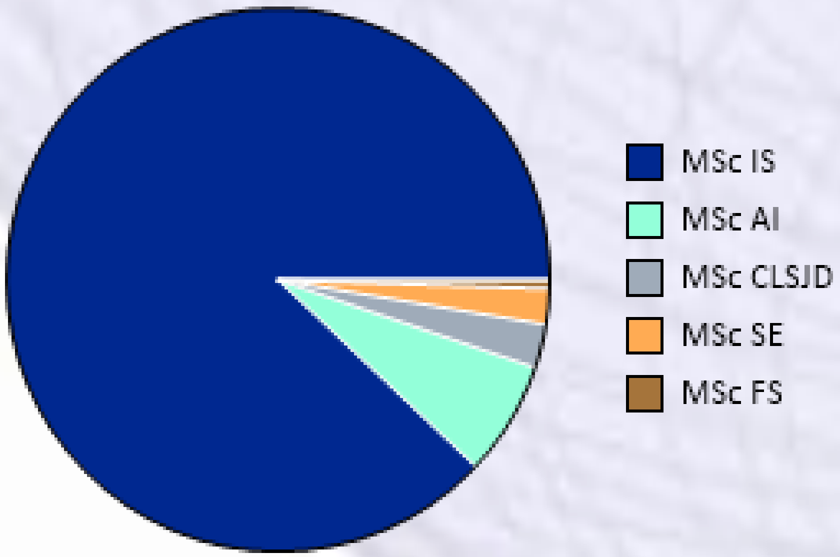
Shi Hu

Sarah Ibrahimi



David Zhang

Students



Admissions forcefully stopped at 173.

Syllabus

Full pdf with course details available on canvas.

Syllabus states course details, deadline, and required reading materials.

The course

				Math tutorial	Lab
1	44	Intro and Regression I	Regression II		Lab Intro
2	45	Classification I	Classification II	Math tutorial 1	**Lab Assignment 1**
3	46	Deep Learning I	Representing Images		**Lab Assignment 2**
4	47	Representing Text	Deep Learning II	Math tutorial 2 **Peer feedback open questions**	Kaggle-Project week 1
5	48	Setting up a ML system *Intro to Reinforcement Learning*	Recommender Systems		Kaggle-Project week 2
6	49	AI for industry and humanities	Autonomous driving	Math tutorial 3	Kaggle-Project week 3
7	50	Q&A			Kaggle-Project week 4
8	**51**	**Exam**			**Poster/Demo Session**

The good, the bad, and the ugly

The bad news

This course is only 8 weeks long.

9 theory lectures

3 expert/applied lectures

1 Q&A session

Note: this course is considered as hard!

The ugly news

The course is packed

Students come from many different backgrounds and experiences.

We are forced by UvA planning to split the practicals in 4 groups.

Each person is randomly assigned to a group.

If your group timeslot is not suitable, you can try and switch with someone yourself. As long as rooms are not overcrowded, I am fine.

The good news

This course is here to teach you a lot.

We are all here to help you out.

Canvas, lectures, labs, tutorials…

Overall schedule

				Math tutorial	**Lab**
1	44	Intro and Regression I	Regression II		Lab Intro
2	45	Classification I	Classification II	Math tutorial 1	**Lab Assignment 1**
3	46	Deep Learning I	Representing Images		**Lab Assignment 2**
4	47	Representing Text	Deep Learning II	Math tutorial 2 **Peer feedback open questions**	Kaggle-Project week 1
5	48	Setting up a ML system *Intro to Reinforcement Learning*	Recommender Systems		Kaggle-Project week 2
6	49	AI for industry and humanities	Autonomous driving	Math tutorial 3	Kaggle-Project week 3
7	50	Q&A			Kaggle-Project week 4
8	**51**	**Exam**			**Poster/Demo Session**

Lab assignments

Three weekly assignment files.

Week 1 has no deliverable, serves as practice.

Deadlines week 2: **10-11-2019 23:59**. Week 3: **17-11-2019 23:59**.

Each week, an ipython notebook to be filled in, with automatic grading!

Tutorials and open questions

Dive deeper into math of machine learning.

Questions partially covered in exam.

Flipped classroom tutorials in week 2, 4, and 6. Come prepared.

Open question assignment (exam practice) on Canvas (due in week 4).

Final project

```
Task: Select the most corresponding reason why this statement is against common sense.
Statement: He put an elephant into the fridge.
A: An elephant is much bigger than a fridge. (correct)
B: Elephants are usually white while fridges are usually white.
C: An elephant cannot eat a fridge.
```

Project 2: Common sense challenge

Project 1: Food recognition challenge

Project 3: Meme analysis challenge

Implement and improve your own full machine learning pipeline!

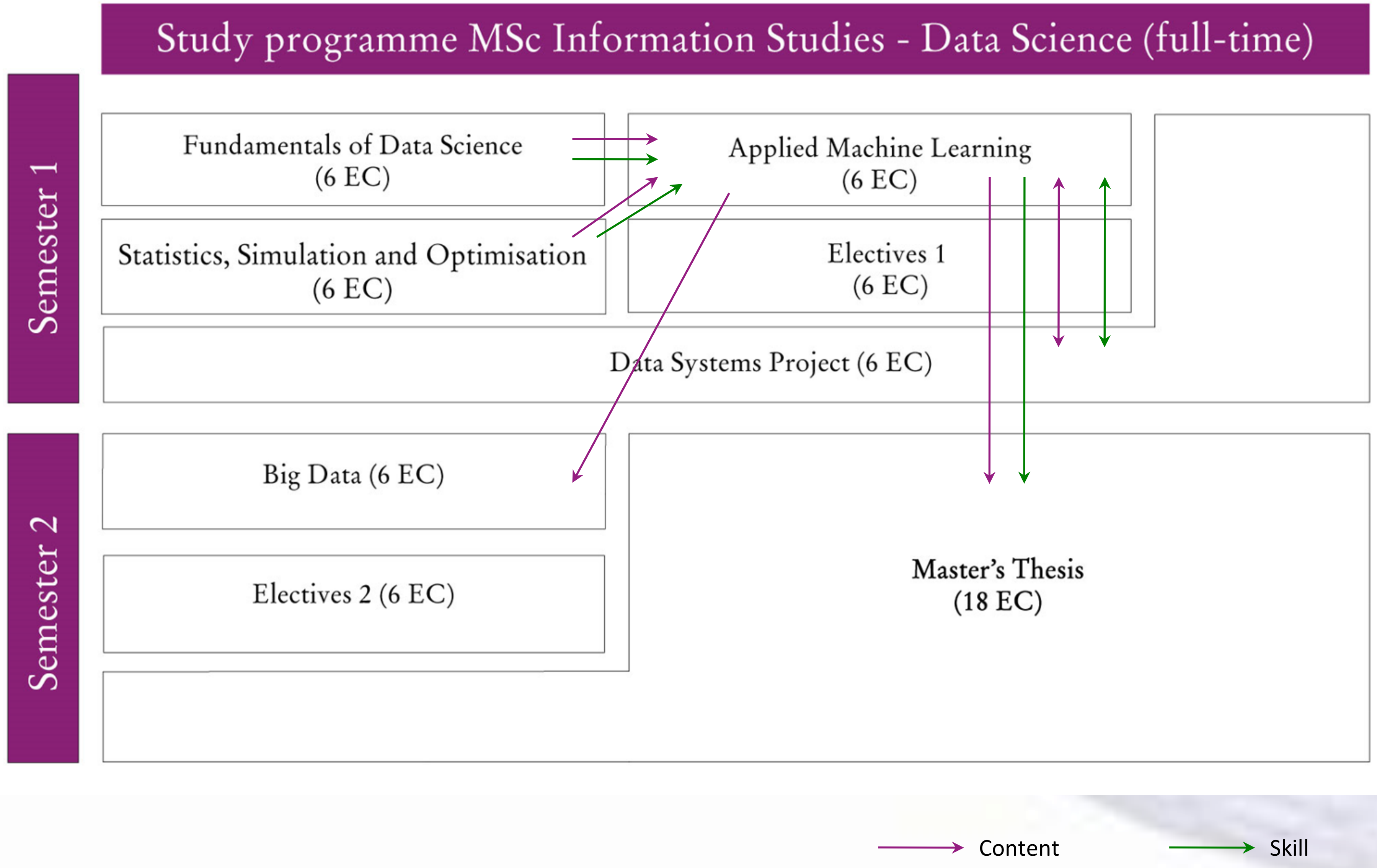
Competition amongst groups of three (groups made on Canvas).

Exam

Consists of closed and open questions.

Covers every lecture.

Examples from previous years on Canvas, naturally without answers.


Study programme MSc Information Studies - Data Science (full-time)
Semester 1
Semester 2
Fundamentals of Data Science (6 EC)
Statistics, Simulation and Optimisation (6 EC)
Applied Machine Learning (6 EC)
Electives 1 (6 EC)
Data Systems Project (6 EC)
Big Data (6 EC)
Electives 2 (6 EC)
Master's Thesis (18 EC)
Content
Skill

Asking questions

For general questions and discussions, we use Canvas and lectures.

Programming and math questions discussed during labs and tutorials.

Questions on grades, groups, etc, mail Cristian:
c.m.rodriguezrivero@uva.nl

Any remaining problem or issues, weekly open office with Pascal:

Every Tuesday from 15:00 to 16:00 in C3.261.

Code of conduct

We encourage you to help each other…

Your grade depends on what you do, not what others do

… and us: give feedback on the course, it can always improve

High expectations

We from you, you from us

Come prepared: read materials, watch videos, look for solutions

Invest time

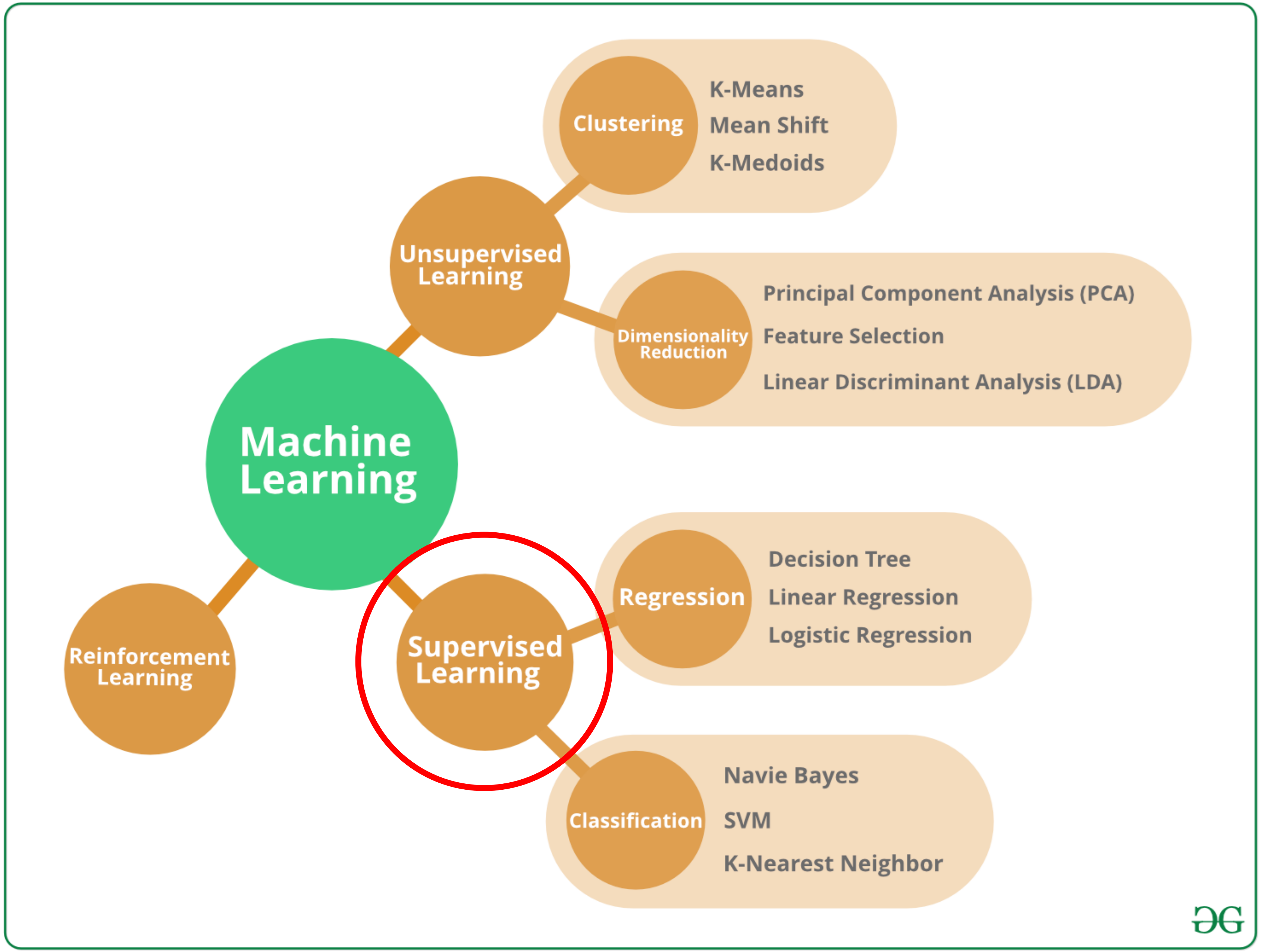
However, we do not tolerate blind copy

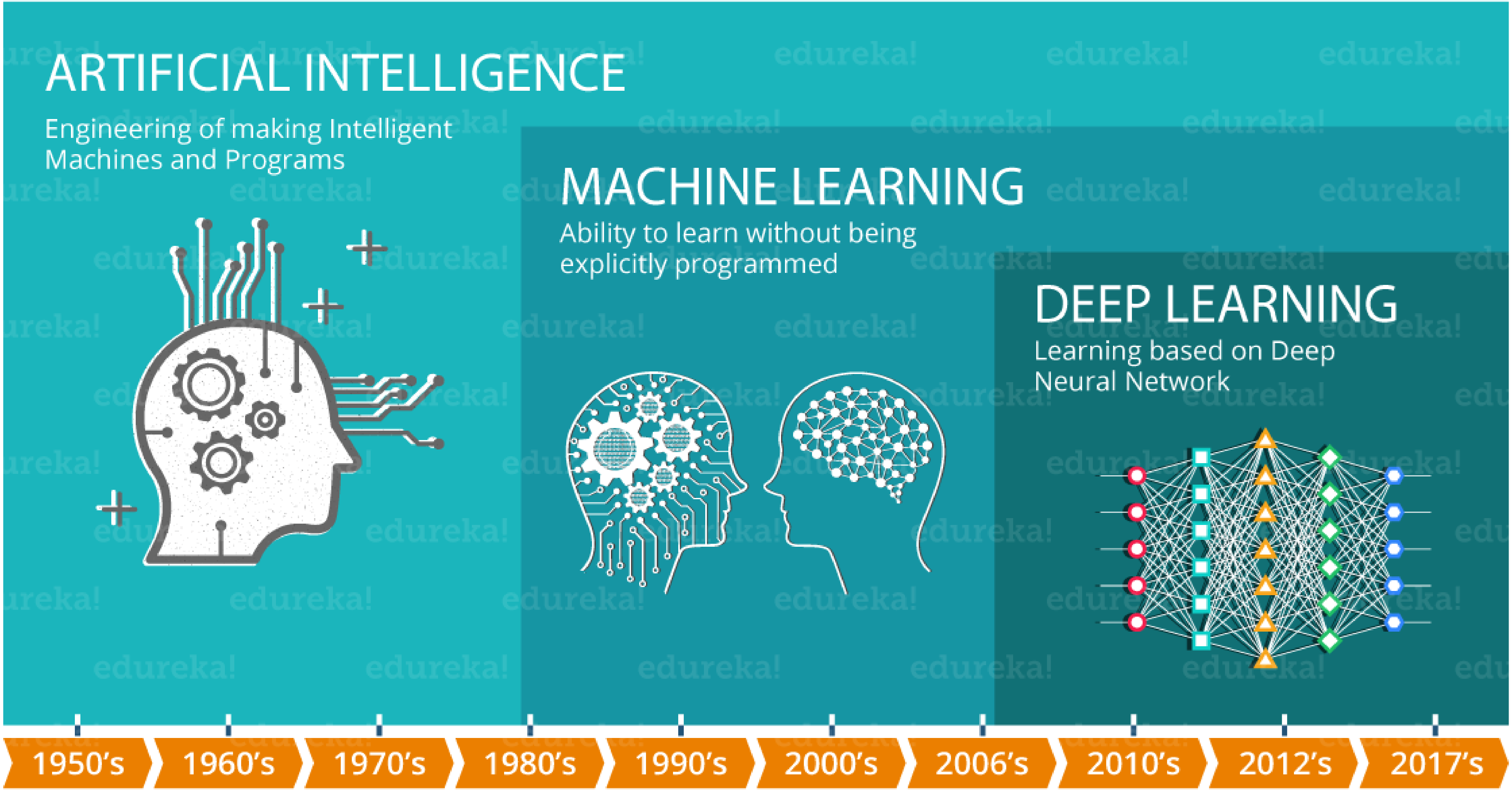
Not from each other

Not from the internet

Applied Machine Learning

Wat is machine learning to you?


K-Means
Clustering
Mean Shift
K-Medoids
Unsupervised Learning
Principal Component Analysis (PCA)
Dimensionality Reduction
Feature Selection
Linear Discriminant Analysis (LDA)
Machine Learning
Decision Tree
Regression
Linear Regression
Logistic Regression
Reinforcement Learning
Supervised Learning
Navie Bayes
Classification
SVM
K-Nearest Neighbor


ARTIFICIAL INTELLIGENCE
Engineering of making Intelligent Machines and Programs
MACHINE LEARNING
Ability to learn without being explicitly programmed
DEEP LEARNING
Learning based on Deep Neural Network
1950's
1960's
1970's
1980's
1990's
2000's
2006's
2010's
2012's
2017's

Machine learning 101

Given input **x**

Do some fancy computing

Predict output **y**

Tom Mitchel, Machine Learning, 1997:

A computer learns, when

for some tasks T; and

performance measure P;

the value P increases for T with experience E.

Types of data

Structured

Sensor readings, tax forms, excel tables

Unstructured

Images, free text, speech

This course will especially focus on unstructured data!

Machine learning lingo

Concepts

Task, system, evaluation, data

Mathematics

Required to formalize and understand the core

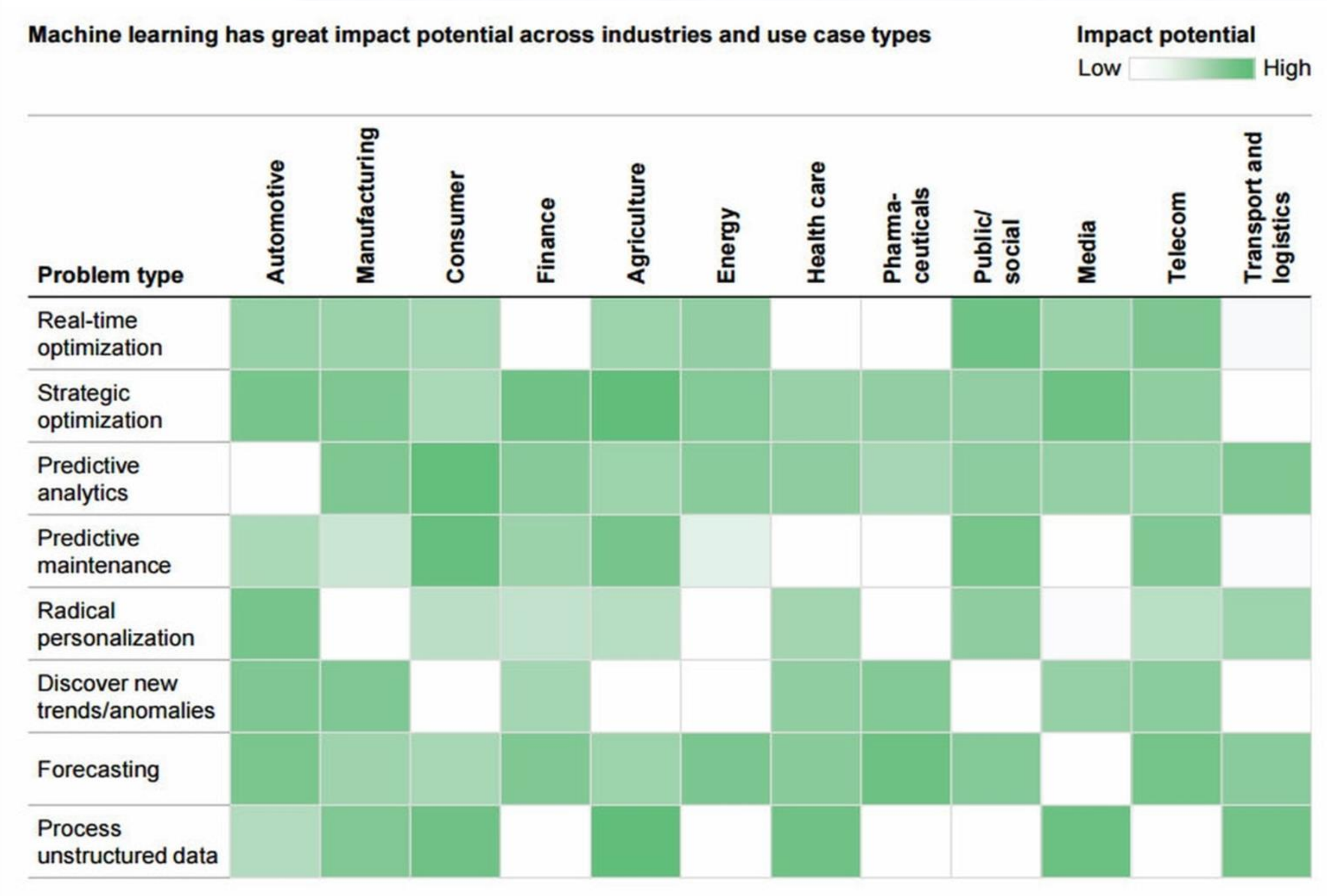
Functions, vectors, matrices, loss, probabilities, predictions

Code

Enable and use machine learning

Python current go to, optionally with sklearn, PyTorch etc.

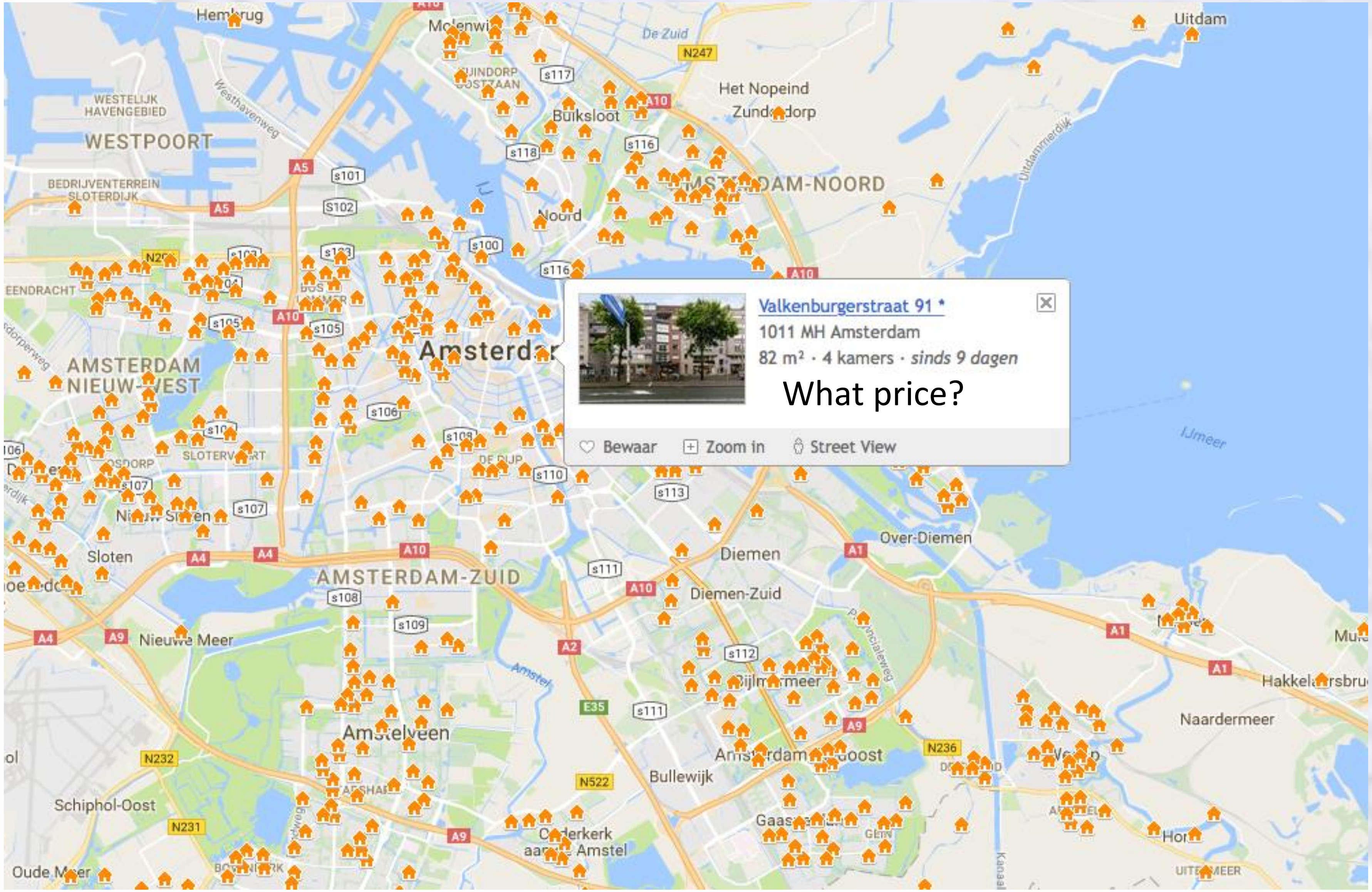
Source: Théo Szymkowiak, “12 sectors where automation will take over in the short term”



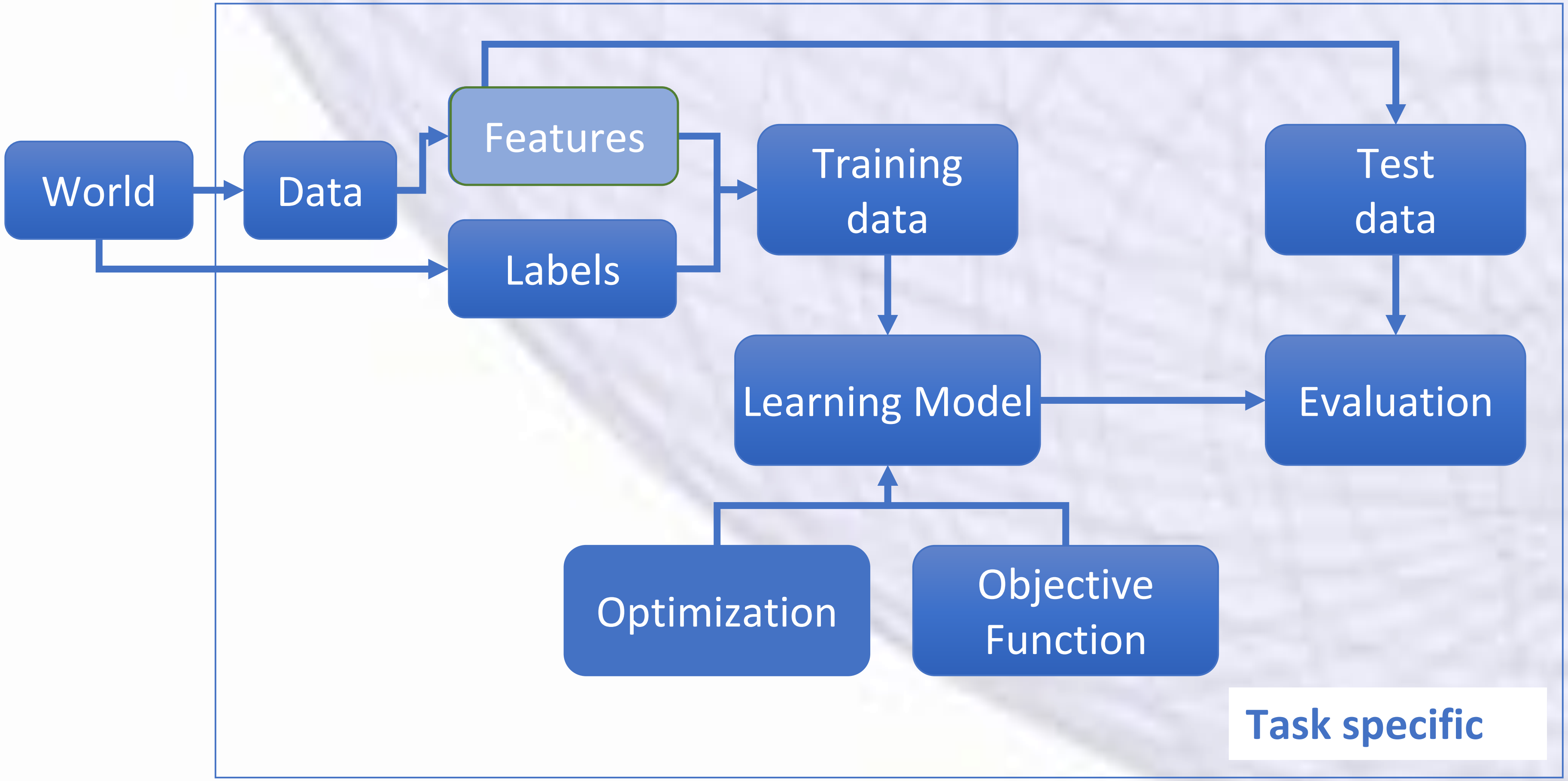
Break

Regression I

Regression


Valkenburgerstraat 91 *
1011 MH Amsterdam
82 m² · 4 kamers · sinds 9 dagen
What price?
Bewaar
Zoom in
Street View

Machine learning system



Input features

Represented as a vector **x**

Representation and values depend on task.

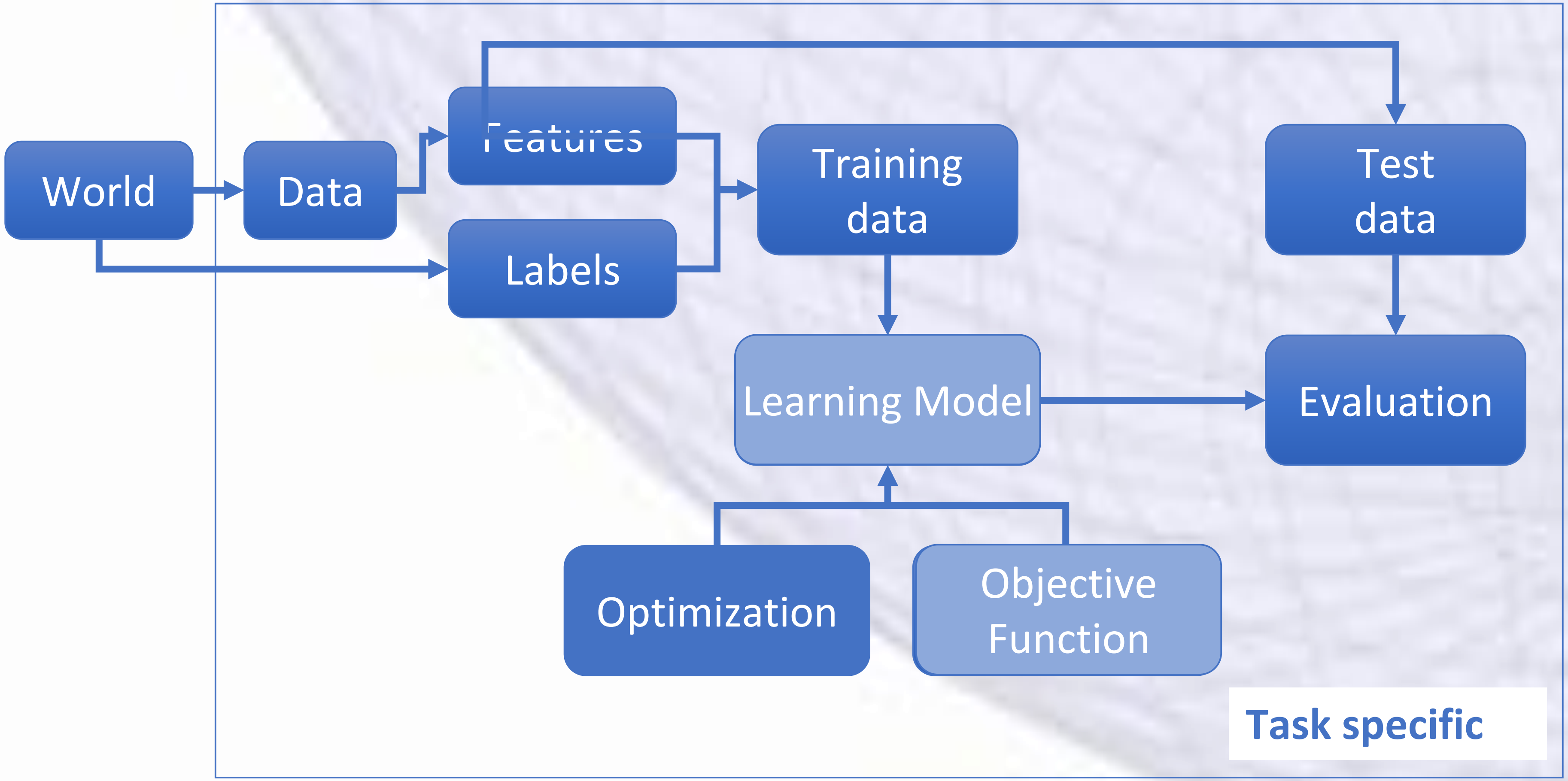
What is a possible feature vector of a house?

E.g. vector [m^2, has_garage, has_balcony, nr_bathrooms, ...]

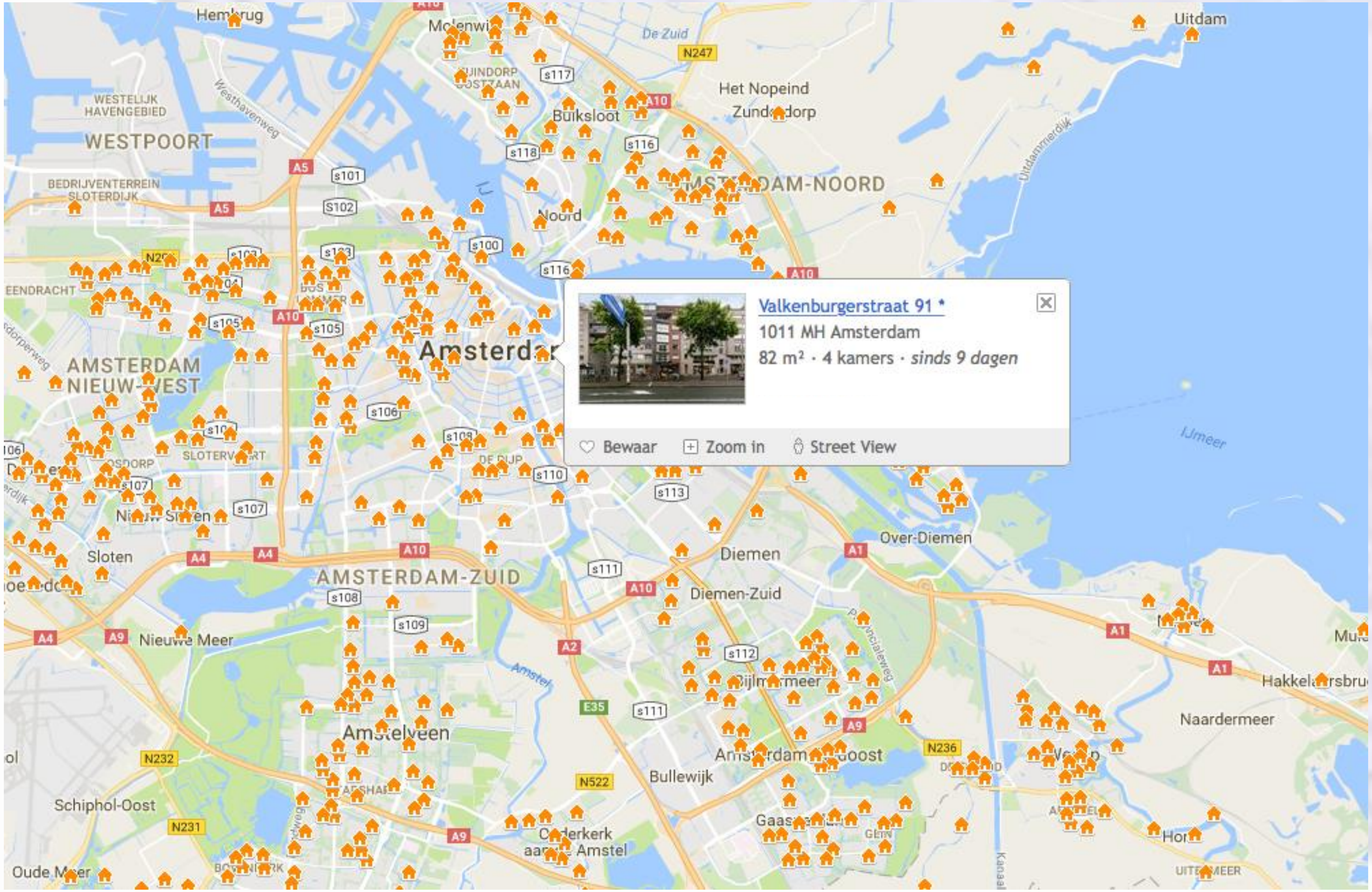
What is a possible feature vector of an image?

Hard to represent, topic of lecture 6!

Machine learning system



Estimating house price



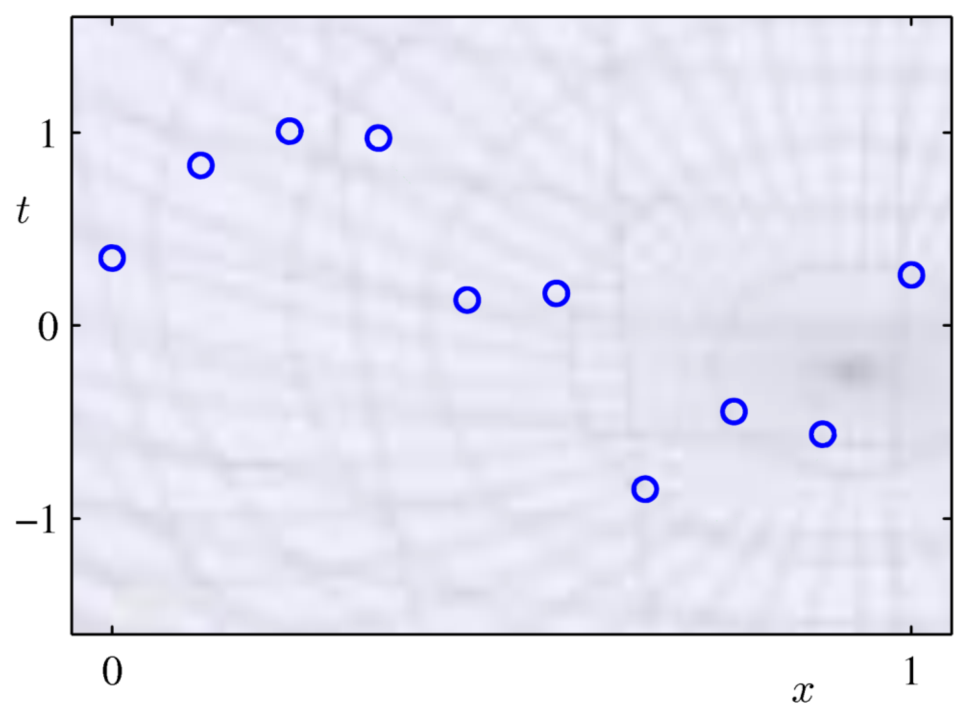
Regression

Predict continuous output value t given input vector **x** (one value here).

$$t = f(\boldsymbol{x}, \boldsymbol{w})$$

The goal is:

- Determine function class f.
- Find best values for **w** from training set.

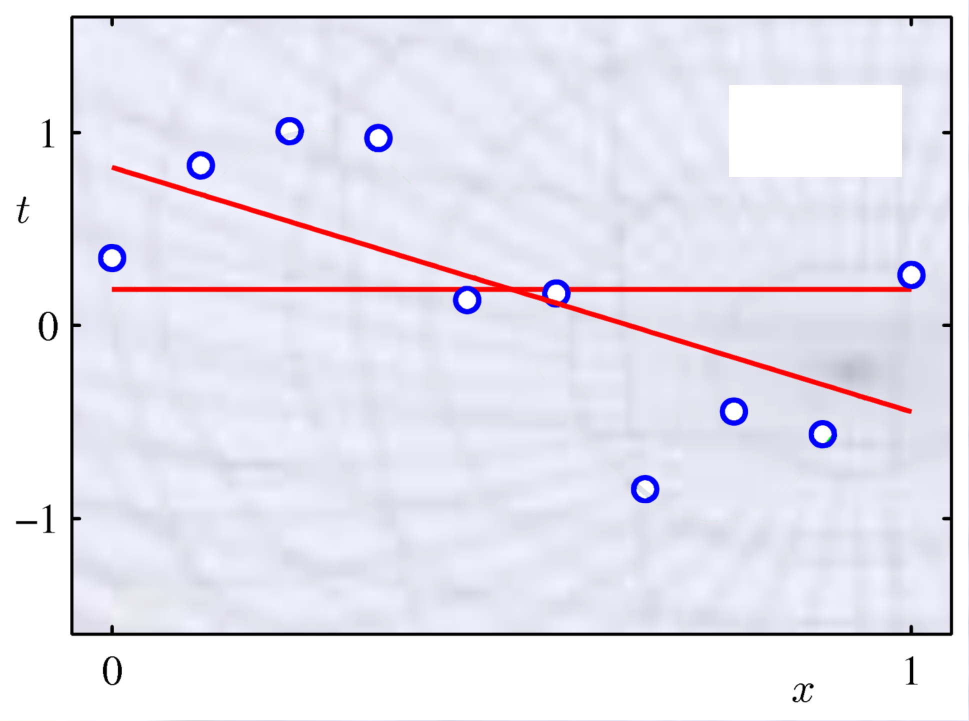


Linear regression

Linear score function:

$$t = f(\boldsymbol{x}, \boldsymbol{w}) = \boldsymbol{x}^T \boldsymbol{w}$$

Number of parameters is equal to the number of features.



Adding an intercept

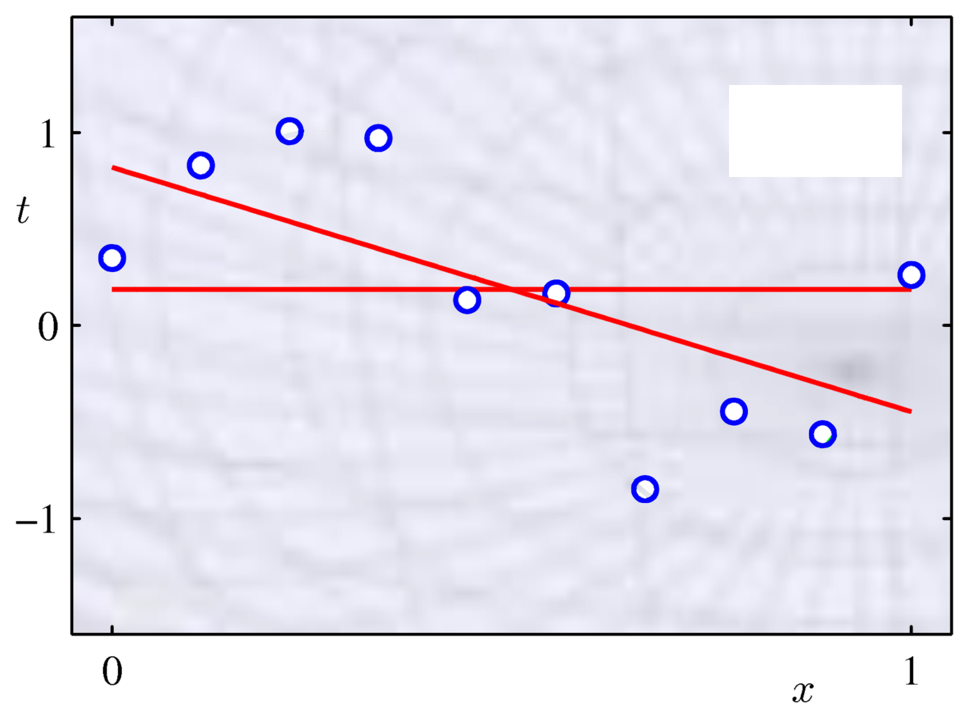
In a standard setup, the output is solely a function of the input.

If $\mathbf{x} = [0,0,..,0]^T$ then the output is always 0.

To overcome this, we add an intercept:

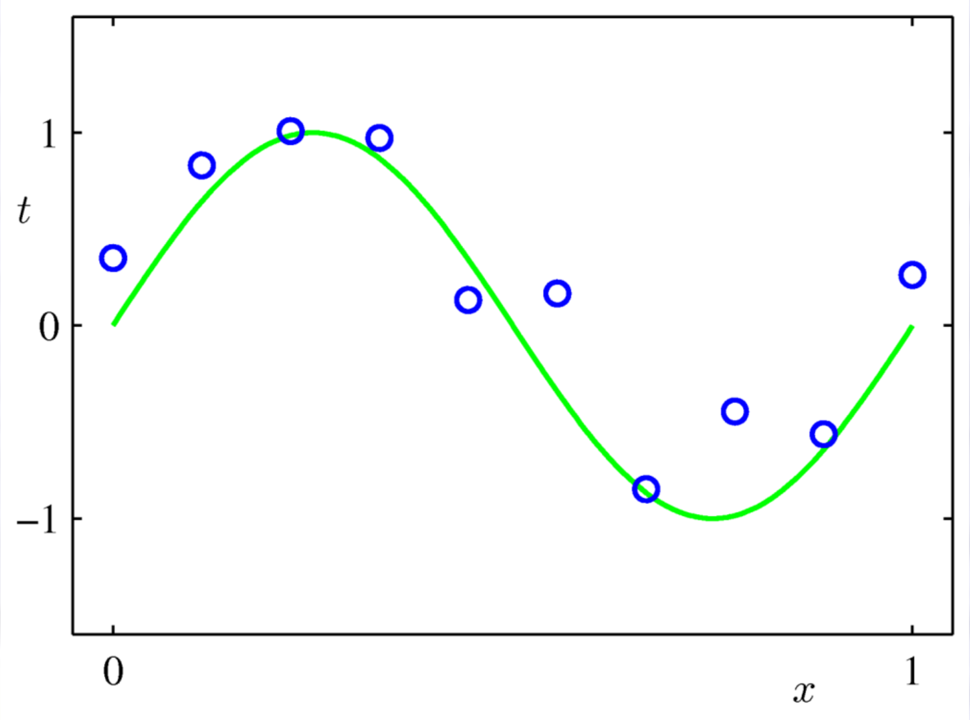
$\mathbf{x} := [\mathbf{x}\ 1]$

$\mathbf{w} := [\mathbf{w}\ b]$

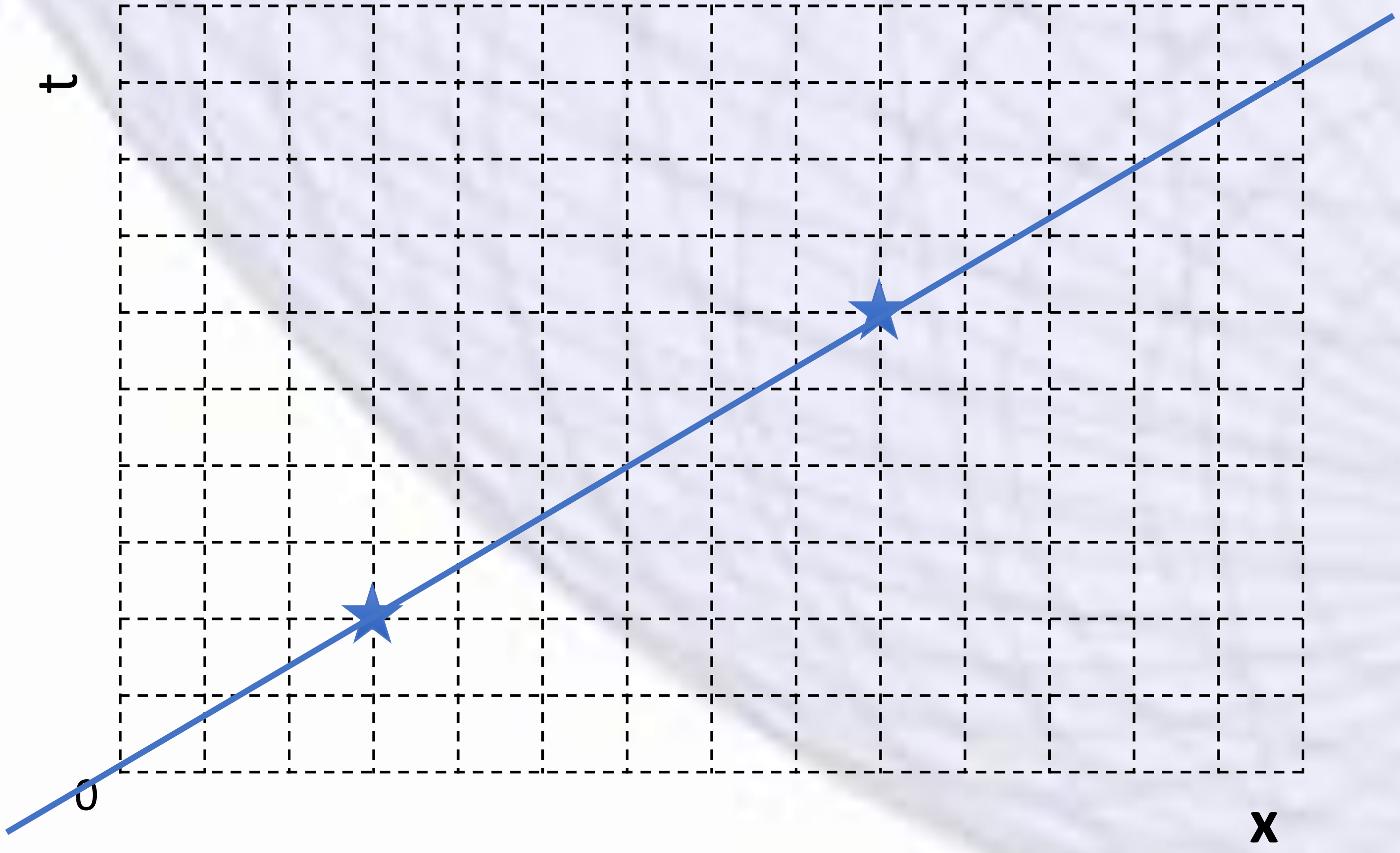


Beyond linear regression

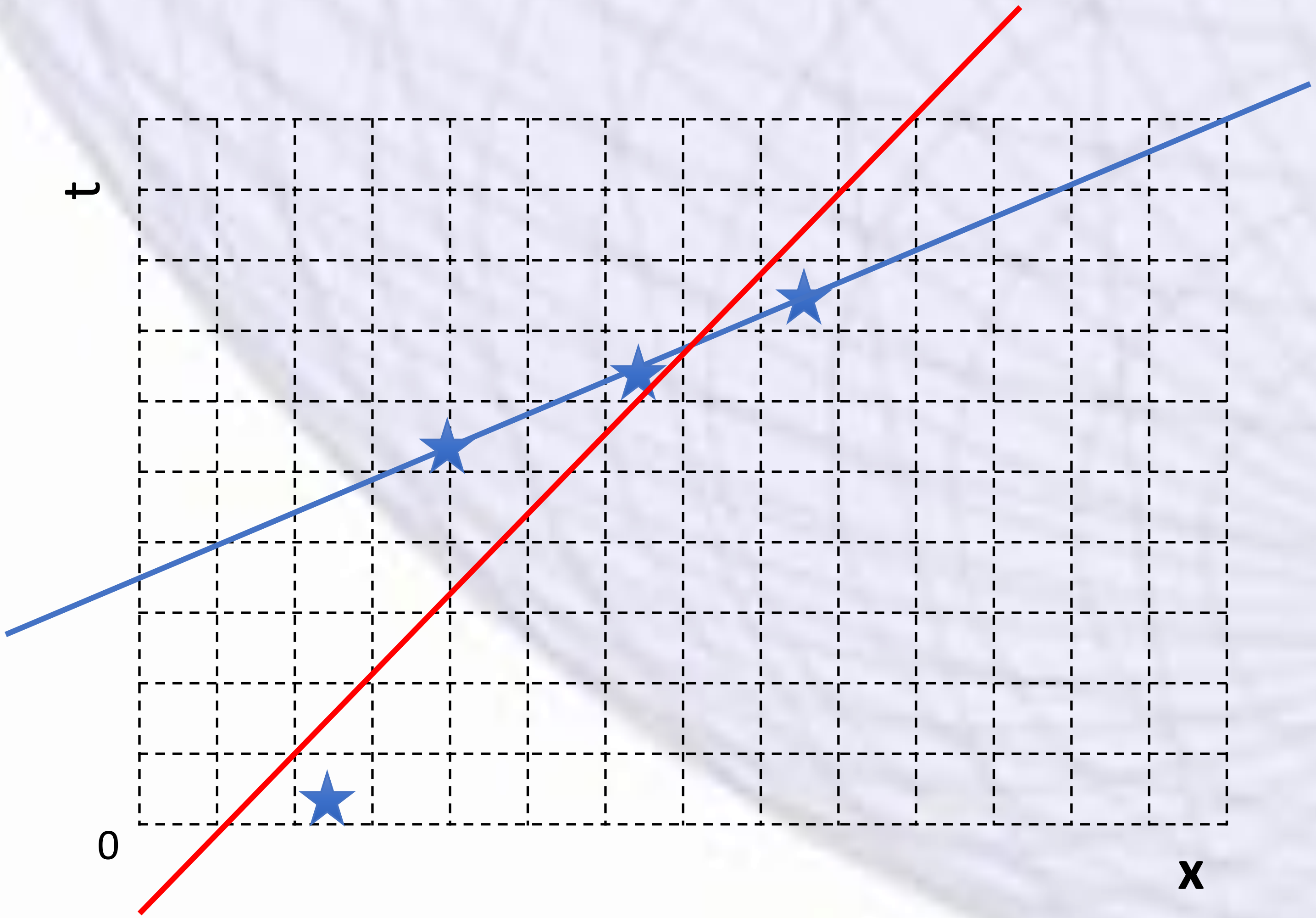
With the proper choice for f, we can do much more than straight lines.



Fitting a line



Fitting a line

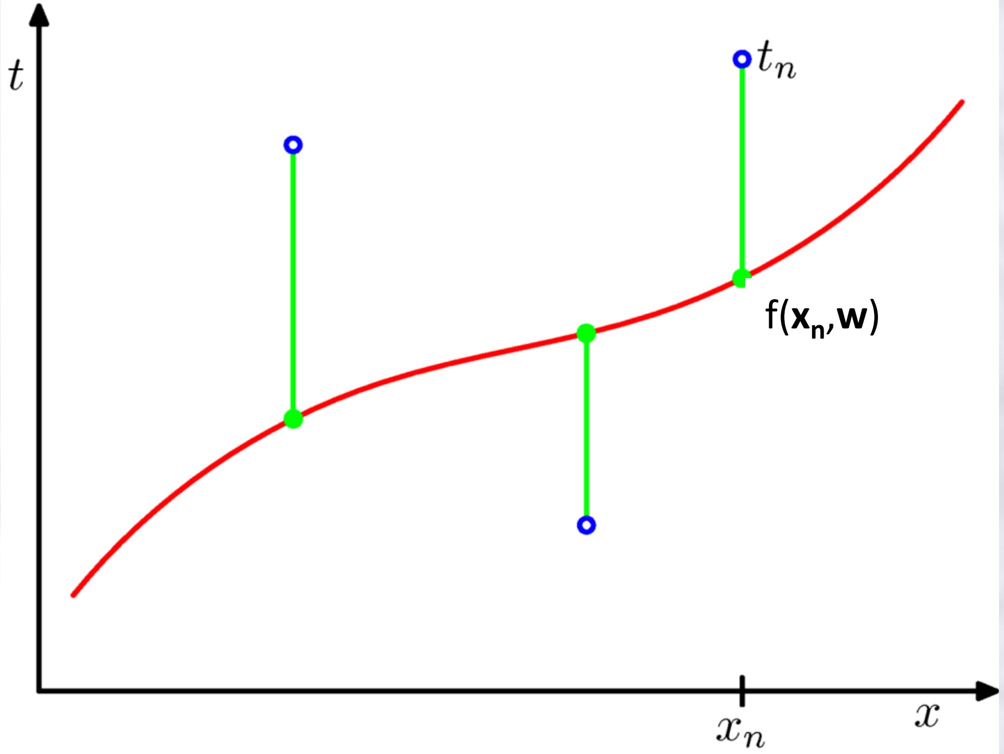


When are predictions good?

Data point x_n

True value t_n

Predicted $\dot{t}_n$



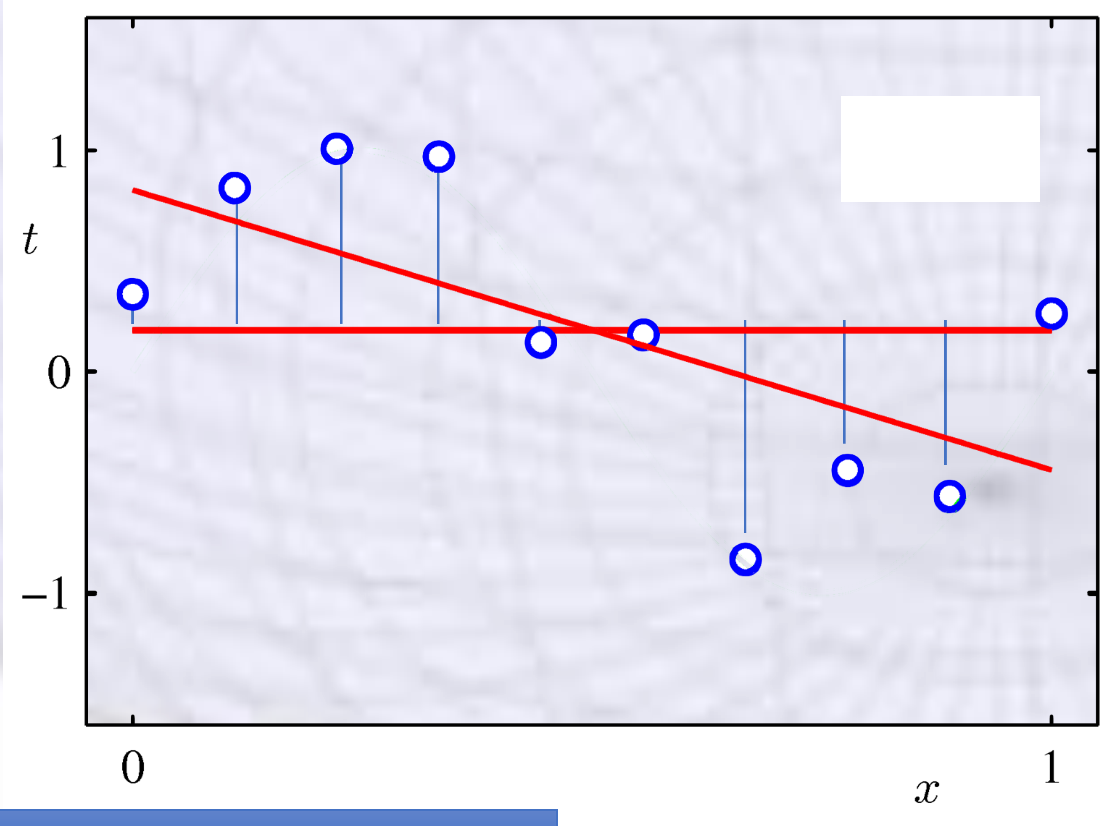
How good is a model?

Loss function: $(t_n - \dot{t}_n)^2$

Function: $f(\boldsymbol{x}, \boldsymbol{w}) = \boldsymbol{x}^T \boldsymbol{w}$

Learning objective:

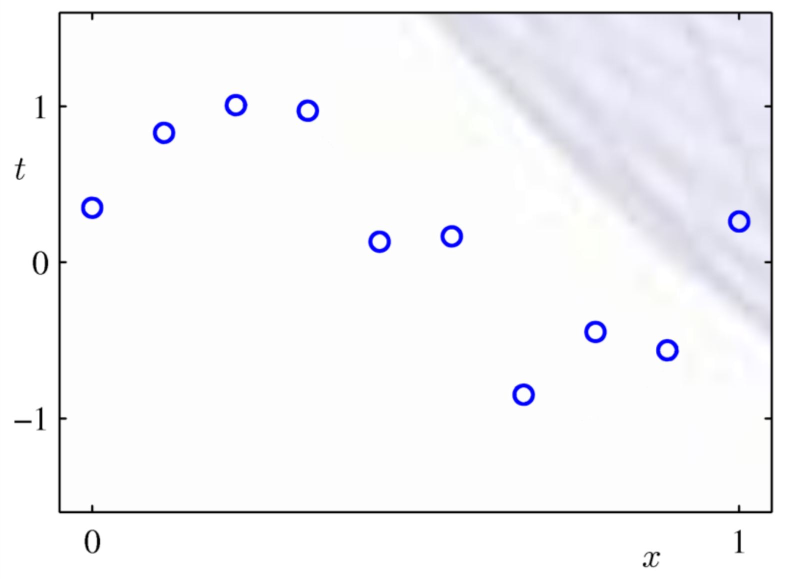
$$s(\boldsymbol{w}) = \frac{1}{2} \sum_n (t_n - \boldsymbol{x}_n{}^T \boldsymbol{w})^2$$



Each model is a line and corresponds to a score, s(w)

What if

I get more data?



I get more features?

x = [a, b, c, d, e, f]

Question: ZIP code feature

You want to add zip code as feature

New representation: x := [x zip] (zip : 1098 ~~XH~~)

It is known by real estate agents that the zipcode is a strong feature, yet your new model performs badly.

1. Explain why this is the case
2. How could you do better?

Content of today

Course logistics.

Machine learning definition and system.

Intro to regression.

What is next

Laptopcollege

Group D tomorrow, other groups on Friday.

Lecture

Regression II, Friday 13:00-15:00, C1.110

Thank you