Machine Learning for Graphics and Vision

Prof. Dr.-Ing. Andreas Geiger

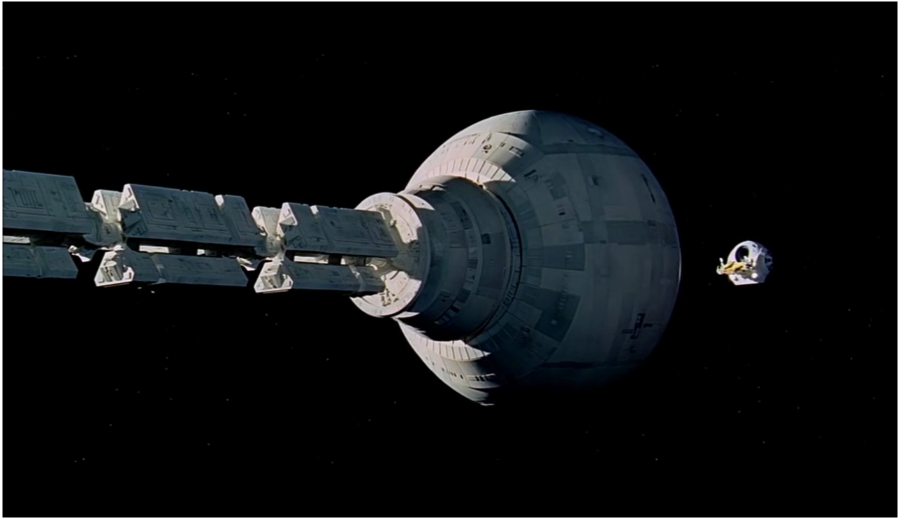
Autonomous Vision Group
MPI-IS / University of Tübingen

April 19, 2018

University of Tübingen
MPI for Intelligent Systems
Autonomous Vision Group

What is Artificial Intelligence?

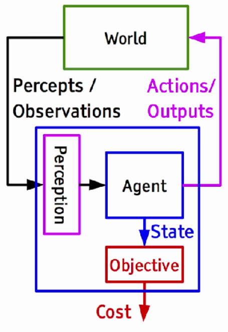
Artificial Intelligence



Artificial Intelligence

"An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves." [John McCarthy]

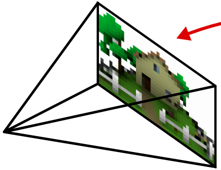
- Machine Learning
- Computer Vision
- Natural Language Processing
- Robotics & Control
- Graphics
- Art, Industry 4.0, Education ...



But: No unique definition (*e.g.*, OCR)

Computer Graphics vs. Computer Vision

2D Image | **3D Scene**



Graphics

Vision

Pixel Matrix

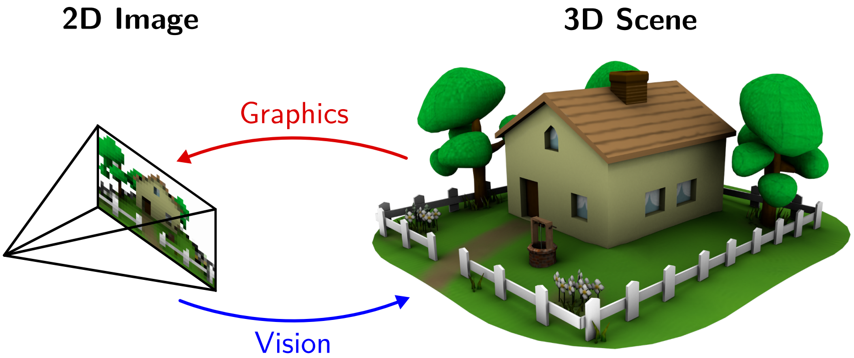
217	191	252	255	239
102	80	200	146	138
159	94	91	121	138
179	106	136	85	41
115	129	83	112	67
94	114	105	111	89

Objects Material

Shape/Geometry Motion

Semantics 3D Pose

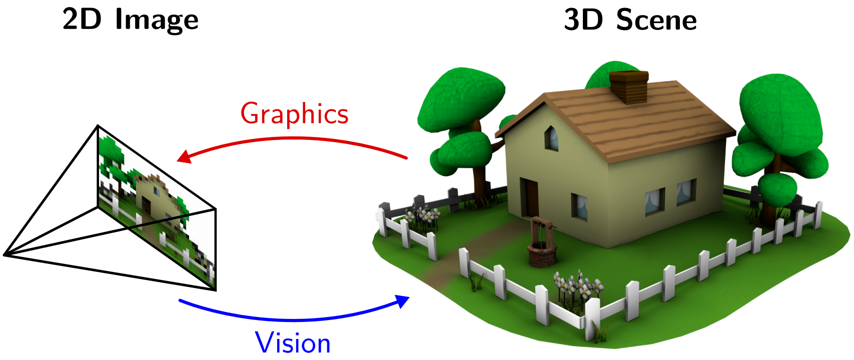
Computer Graphics vs. Computer Vision



Computer Vision is an ill-posed inverse problem:

- Many 3D scenes yield the same 2D image
- Additional constraints (knowledge about world) required

Computer Graphics vs. Computer Vision



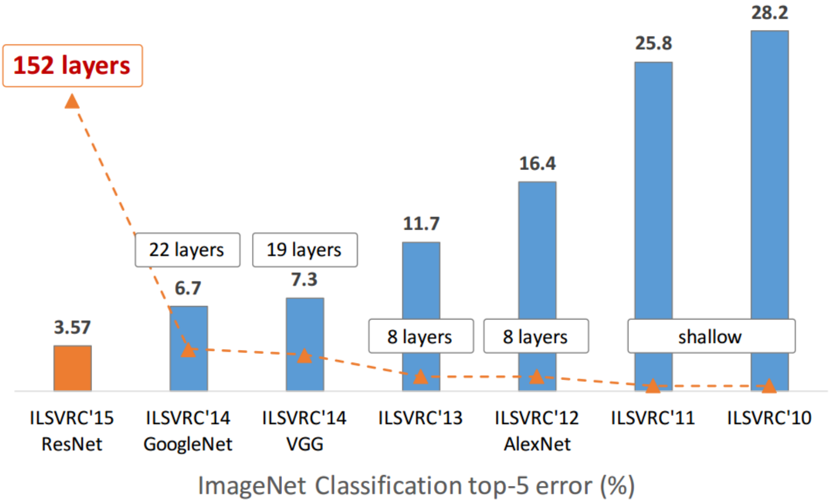
Computer Graphics can also be an ill-posed problem:

- Rendering is slow, fill-in missing information
- Deblurring, denoising, superresolution

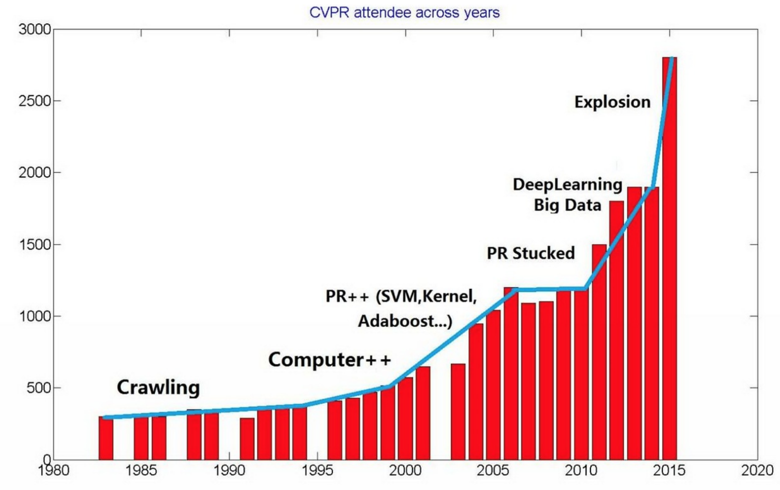
What is all the fuss about?

www.image-net.org

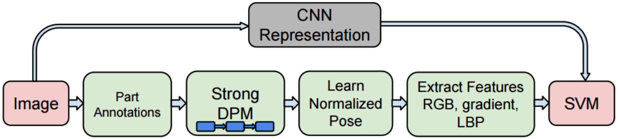
The Deep Learning Revolution

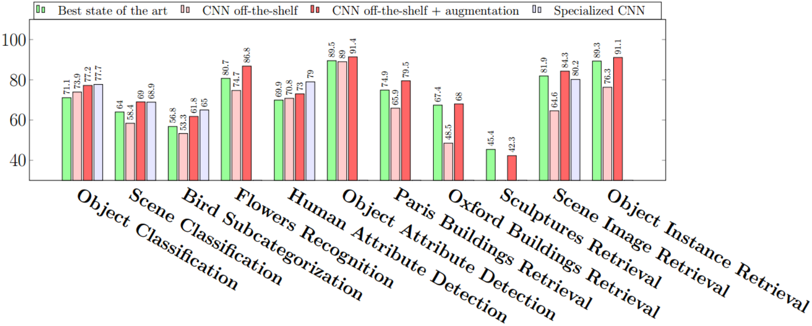


CVPR Attendance



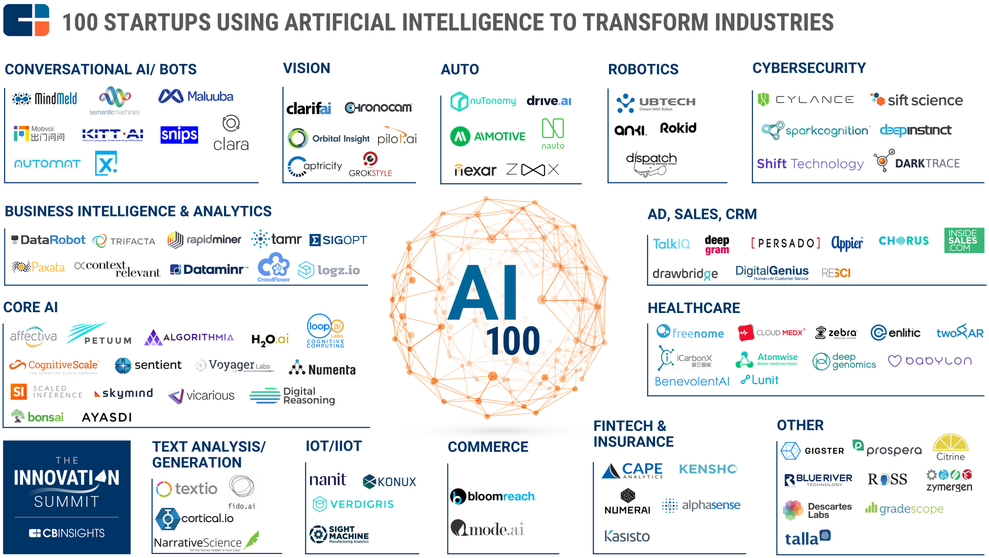
Generalization





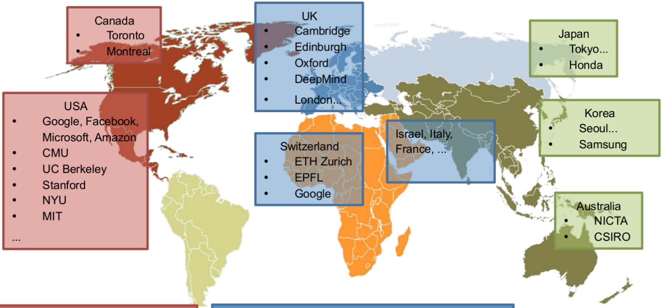
[Razavian, Arxiv 2014]

AI Startups



100 STARTUPS USING ARTIFICIAL INTELLIGENCE TO TRANSFORM INDUSTRIES
CONVERSATIONAL AI/ BOTS
MindMeld
semanticmachines
Maluuba
Mobvoi 出门问问
KITT.AI
snips
clara
AUTOMAT
X.
VISION
clarifai
Chronocam
Orbital Insight
pilot.ai
Captricity
GROKSTYLE
AUTO
nuTonomy
drive.ai
AIMOTIVE
nauto
nexar
ZOOX
ROBOTICS
UBTECH
Dream With Robot
anki
Rokid
dispatch
CYBERSECURITY
CYLANCE
sift science
sparkcognition
deepinstinct
Shift Technology
DARKTRACE
BUSINESS INTELLIGENCE & ANALYTICS
DataRobot
TRIFACTA
rapidminer
tamr
SIGOPT
Paxata
context relevant
Dataminr
CrowdFlower
logz.io
AD, SALES, CRM
TalkIQ
deep gram
[PERSADO]
Appier
CHORUS
INSIDE SALES .COM
drawbridge
DigitalGenius
Human+AI Customer Service
RESCI
AI
100
CORE AI
affectiva
PETUUM
ALGORITHMIA
H2O.ai
loop ai Labs
COGNITIVE COMPUTING
CognitiveScale
THE COGNITIVE CLOUD COMPANY
sentient
Voyager Labs
Numenta
SI SCALED INFERENCE
skymind
vicarious
Digital Reasoning
bonsai
AYASDI
HEALTHCARE
freenome
CLOUD MEDX
zebra
enlitic
twoXAR
iCarbonX
Atomwise
Better medicines faster.
deep genomics
babylon
BenevolentAI
Lunit
THE INNOVATION SUMMIT
CBINSIGHTS
TEXT ANALYSIS/ GENERATION
textio
fido.ai
cortical.io
NarrativeScience
IOT/IIOT
nanit
KONUX
VERDIGRIS
SIGHT MACHINE
Manufacturing Analytics
COMMERCE
bloomreach
mode.ai
FINTECH & INSURANCE
CAPE ANALYTICS
KENSHO
NUMERAI
alphasense
Kasisto
OTHER
GIGSTER
prospera
Citrine
BLUE RIVER TECHNOLOGY
ROSS
zymergen
Descartes Labs
gradescope
talla

AI Research Worldwide



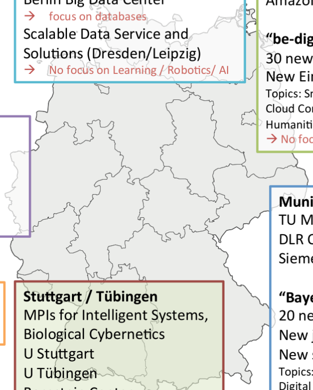
Facebook AI Research
- New lab founded in 2013
- New arm in Paris: 2015
- Europe University partnership program 02/2016

Google DeepMind (London)
- Founded in 2011 as AI research startup
- 2014 acquisition through Google ($0.4bn)
- 02/2016: DeepMind learning algorithm beats European Go champion

Toyota Research Institue (Palo Alto)
- 11/2015 Toyota announced that it is investing $1bn in the next 5 years
- New AI and robotics R&D arm in Silicon Valley
- Two research centers at Stanford and MIT ($50m in 5 years)

OpenAI (San Francisco)
- Non profit AI research company founded in 12/2015
- $1bn pledged from Elon Musk, Amazon and others

AI Research in Germany



BMBF Centers
Berlin Big Data Center
→ focus on databases
Scalable Data Service and Solutions (Dresden/Leipzig)
→ No focus on Learning / Robotics/ AI

Berlin
TU Berlin
Amazon Research

"be-digital.berlin" (announced 12/2015)
30 new university chairs in IT
New Einstein-Center
Topics: Smart Cities, Industry 4.0, Big Data, Cloud Computing, Smart Mobility, Digital Humanities
→ No focus on Learning / Robotics/ AI

Saarbrücken / Kaiserslautern
MPIs Computer Science, Software Systems

DFKI
→ Applied research

Munich
TU Munich, LMU
DLR Oberpfaffenhofen
Siemens, BMW, ...

"Bayern Digital" (announced 07/2015)
20 new university chairs
New junior groups / PhD program
New startup centers / WERK1.Bayern
Topics: Industry 4.0, Networks and Mobility, Digital Health, Smart Energy, Cyber Security
→ No focus on Learning / Robotics / AI. Broad.

Karlsruhe
KIT
→ Robotics

Stuttgart / Tübingen
MPIs for Intelligent Systems, Biological Cybernetics
U Stuttgart
U Tübingen
Bernstein Center
Center for Integrative Neurosciences
Bosch, Daimler, Trumpf, etc.

Augsburg
KUKA
→ Robotics

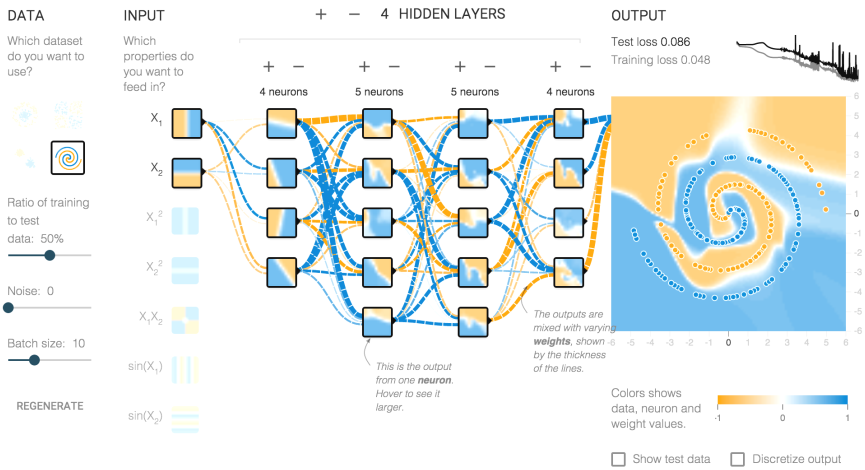
Cyber Valley

- Internationally visible hub for research on AI & IS
- Max Planck, universities, government and industry
- Foster entrepreneurship, educate engineers for the IS economy, IMPRS

Why is the AI revolution happening now?

#1 Machine Learning Starts Working

Machine Learning Starts Working



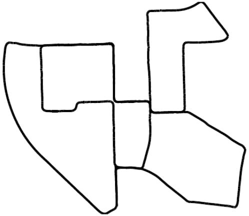
playground.tensorflow.org

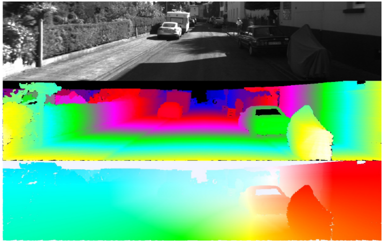
#2 Big Data

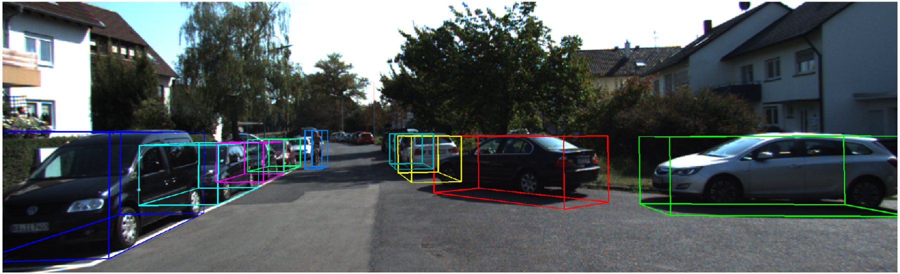
Big Data

www.image-net.org

Big Data







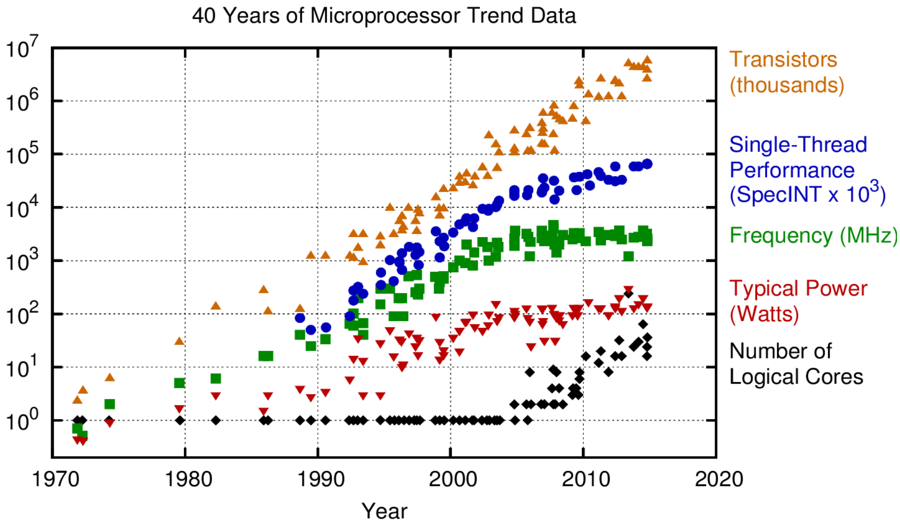
www.cvlibs.net/datasets/kitti/

Big Data

www.cityscapes-dataset.com

#3 Moore’s Law

Moore’s Law

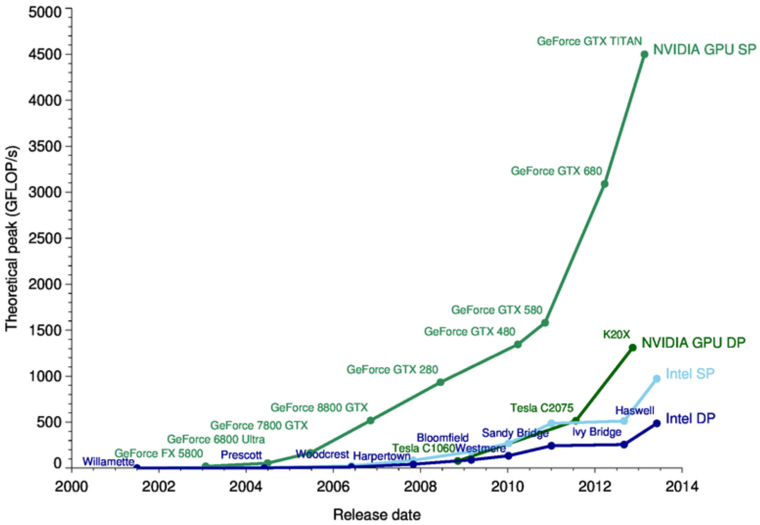


Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2015 by K. Rupp

Moore's Law

www.nvidia.com

Moore’s Law



Theoretical peak (GFLOP/s)
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
Release date
2000
2002
2004
2006
2008
2010
2012
2014
GeForce GTX TITAN
NVIDIA GPU SP
GeForce GTX 680
GeForce GTX 580
GeForce GTX 480
K20X
NVIDIA GPU DP
GeForce GTX 280
Intel SP
GeForce 8800 GTX
Tesla C2075
Haswell
Intel DP
GeForce 7800 GTX
Bloomfield
Sandy Bridge
Ivy Bridge
GeForce 6800 Ultra
Tesla C1060
Westmere
GeForce FX 5800
Prescott
Woodcrest
Harpertown
Willamette

#4 Coorporate Research Philosophy

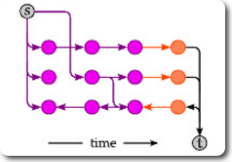
Coorporate Research Philosophy

- Geoffrey Hinton $\rightarrow$ Google
- Yann LeCun $\rightarrow$ Facebook
- Andrew Ng $\rightarrow$ Baidu
- Zoubin Ghahramani $\rightarrow$ Uber
- Ian Goodfellow $\rightarrow$ OpenAI
- Marc Pollefeys $\rightarrow$ Microsoft
- Trevor Darrell $\rightarrow$ Nexar
- Fei-Fei Li $\rightarrow$ Google
- Marc'Aurelio Ranzato $\rightarrow$ Google
- Alex Smola $\rightarrow$ Amazon
- David Nister $\rightarrow$ Tesla
- Raquel Urtasun$\rightarrow$ Uber
- Pushmeet Kohli $\rightarrow$ Google

#5 Open Access

Open Access

3D Reconstruction
Probabilistic volumetric 3D reconstruction with ray potentials.



FollowMe
Computationally Efficient Online Min-Cost Flow Tracking.

Discrete Flow
Optical flow estimation cast as inference in a discrete CRF.

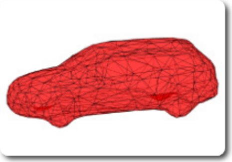
$$f(\mathbf{x}) = \begin{cases} \xi(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases}$$

$$\xi(\mathbf{x}) : D^N \to \mathbb{R}$$

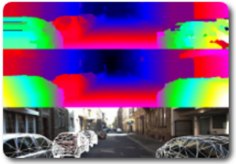
$$|\mathcal{S}| \ll 2^N$$

LIBSMS
Efficient MAP inference with sparse high-order potentials.

Indoor Scenes
3D layout and 3D object inference from a single RGB-D image.



Semi-Convex Hull
Reduces a complex CAD model to a watertight low-face mesh.



Displets
Resolving stereo ambiguities using knowledge about objects.

Object Scene Flow
Scene flow for autonomous vehicles. Novel realistic dataset.

www.cvlibs.net

Open Access

ubuntu

boost
C++ LIBRARIES

python™

TensorFlow™

scikits learn
machine learning in Python

Google SUMMER of CODE 2015

android

Research at Google

protobuf
Protocol Buffers

chromium

Ceres Solver

The KITTI Vision Benchmark Suite
A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago

DAIMLER
CITYSCAPES DATASET

Mapillary

facebook.

torch

Caffe2

PYTORCH

Microsoft Research KINECT

Qt The Qt Company

NVIDIA CUDA

id

DOOM

OpenCV

Willow Garage

ROS

pcl

OpenStreetMap
The Free Wiki World Map

#6 Rapid Dissemination

Rapid Dissemination

Cornell University Library

We gratefully acknowledge support from the Simons Foundation and The Alliance of Science Organisations in Germany, coordinated by TIB, MPG and HGF

arXiv.org > cs > arXiv:1701.04722

Search or Article ID inside arXiv | All papers | Broaden your search using Semantic Scholar

(Help | Advanced search)

Computer Science > Learning

Adversarial Variational Bayes: Unifying Variational Autoencoders and Generative Adversarial Networks

Lars Mescheder, Sebastian Nowozin, Andreas Geiger

(Submitted on 17 Jan 2017 (v1), last revised 8 Mar 2017 (this version, v2))

Variational Autoencoders (VAEs) are expressive latent variable models that can be used to learn complex probability distributions from training data. However, the quality of the resulting model crucially relies on the expressiveness of the inference model. We introduce Adversarial Variational Bayes (AVB), a technique for training Variational Autoencoders with arbitrarily expressive inference models. We achieve this by introducing an auxiliary discriminative network that allows to rephrase the maximum-likelihood-problem as a two-player game, hence establishing a principled connection between VAEs and Generative Adversarial Networks (GANs). We show that in the nonparametric limit our method yields an exact maximum-likelihood assignment for the parameters of the generative model, as well as the exact posterior distribution over the latent variables given an observation. Contrary to competing approaches which combine VAEs with GANs, our approach has a clear theoretical justification, retains most advantages of standard Variational Autoencoders and is easy to implement.

Subjects: **Learning (cs.LG)**
Cite as: arXiv:1701.04722 **[cs.LG]**
(or **arXiv:1701.04722v2 [cs.LG]** for this version)

Submission history

From: Lars Mescheder [view email]
[v1] Tue, 17 Jan 2017 15:18:31 GMT (2048kb,D)
[v2] Wed, 8 Mar 2017 13:46:01 GMT (3115kb,D)

Download:

- PDF
- Other formats

(license)

Current browse context:
cs.LG
< prev | next >
new | recent | 1701

Change to browse by:
cs

References & Citations

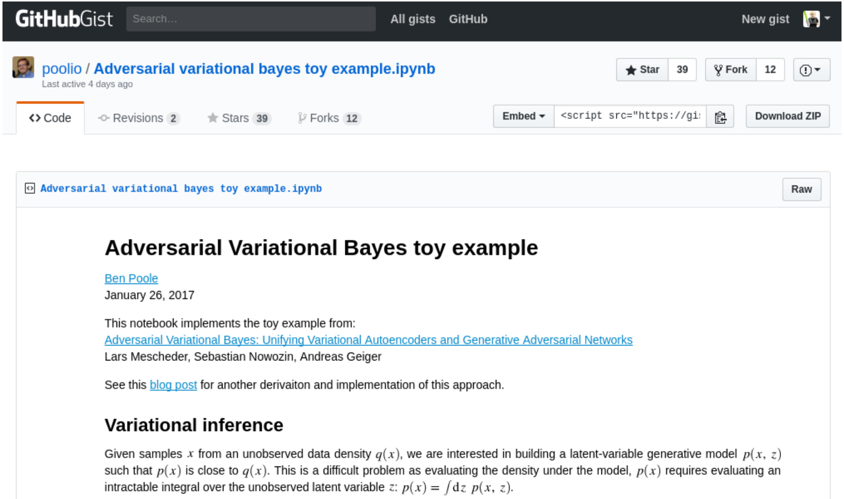
- NASA ADS

DBLP - CS Bibliography
listing | bibtex
Lars M. Mescheder
Sebastian Nowozin
Andreas Geiger

Bookmark (what is this?)

[Mescheder *et al.*, Arxiv 2017]

Rapid Dissemination



`gist.github.com/poolio/b71eb943d6537d01f46e7b20e9225149`

Machine Learning for Graphics and Vision

Team

Prof. Dr. Hendrik Lensch

Computer Graphics Group

Prof. Dr. Andreas Geiger

Autonomous Vision Group

- TAs:
  - CGG: Patrick Wieschollek, Arijit Mallick
  - AVG: Joel Janai, Despoina Paschalidou, Lars Mescheder, Aseem Behl

Contents

Goal: Obtain understanding for basic machine learning concepts and their application to computer vision and computer graphics problems.

Machine Learning Topics:

- Classification / Regression
- kNN, SVM, Random Forest
- Deep Neural Networks (CNN, RNN)
- Latent Variable Models (PCA, VAE)
- Generative Models (GANs)
- Markov Random Fields

Vision and Graphics Applications:

- Object Recognition
- Semantic Segmentation
- Reconstruction & Optical Flow
- Deblurring, Denoising, Superresolution
- Rendering, Light Transport
- Global Illumination Sampling

Organization

- SWS: 2 V + 2 Ü, 6 ECTS
- Lecture: Thursdays, 8:15-10:00, F122 (Kleiner Hörsaal), starting April 19
- Exercise: Fridays, 8:15-10:00, F122 (Kleiner Hörsaal), starting April 20
- Lectures and exercises will be held in English
- Oral exam
  - Dates: 31.7.2018 and 18.9.2018 (mark your calendar!)
  - 0.3 bonus if 60 % of the points in the exercises
- Course Website:

  `www.wsi.uni-tuebingen.de/lehrstuehle/computergrafik/lehrstuhl/teaching/vorlesung-machine-learning-in-graphics-vision.html`
- **Enroll via ILIAS!**

Exercises

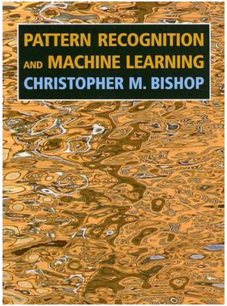
- Every lecture followed by an exercise
  - 3 Tutorials: Python, Tensorflow I, Tensorflow II
  - 6 Assignments (managed via ILIAS, see website), 20 points per exercise
- Purpose: Introduction to assignments, Q&A, discussion of results
- Approx. 12 days for completing each assignment
  - Handout on Friday; Submission until Wednesday 9pm
- Assignments can be done in groups of up to 2 students
  - Establish groups in first exercise session (tomorrow!)
  - Solutions of a group must indicate all group members
  - Every group member must submit the solution
- Solutions may not be shared across groups
- Accessing the computer pool: see course website for instructions

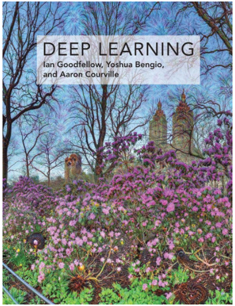
Lectures

19.04.	01	Machine Learning I
26.04.	02	Machine Learning II
03.05.	03	Deep Learning
17.05.	04	Convolutional Neural Networks
07.06.	05	Advanced Topics in Deep Learning
14.06.	06	Reconstruction & Motion Estimation
21.06.	07	Computer Graphics
28.06.	08	Latent Variable Models
05.07.	09	Generative Models
12.07.	10	Structured Prediction I
19.07.	11	Structured Prediction II
26.07.	12	Deep Structured Models

Materials

- Lecture slides, class-room writing (blackboard)
- Exercise slides & assignments
- Books:
  - Bishop: Pattern Recognition and Machine Learning (Springer)
  - Goodfellow, Bengio, Courville: Deep Learning (`www.deeplearningbook.org/`)





Questions?

Introduction

Machine Learning

“Field of study that gives computers ability to learn without explicit programming.”

Arthur Samuel, 1959

Machine Learning

”Predicting the future based on past observations.”

Machine Learning

Für größere Ansicht Maus über das Bild ziehen

Asus F540LA-XX274T 39,6 cm (15,6 Zoll) Notebook (Intel core i3-5005U, 8GB Arbeitsspeicher, 1TB Festplatte, Intel HD 5500, Win 10 Home) schwarz

von Asus Computer

16 Kundenrezensionen | 38 beantwortete Fragen

Erhältlich bei diesen Anbietern.

Kunden, die diesen Artikel gekauft haben, kauften auch

Microsoft Office 365 Home 5PCs/MACs - 1 Jahresabonnement...
Microsoft
47
Windows 7 / 8 / 10
EUR 73,99 Prime

AVM FRITZ!Box 7490 WLAN AC + N Router (VDSL/ADSL, 1.300 Mbit/s (5 GHz), 450 Mbit/s (2,4...
2.460
EUR 193,90 Prime

Tonor 2.4GHz kabellos schnurlos Funkmaus Wireless Optische Maus mit USB 2.0 Empfänger...
24
EUR 8,99 Prime

Logitech MK710 Wireless Desktop Combo Tastatur und Maus (QWERTZ, deutsches Tastaturlayout)
303
EUR 73,99 Prime

Machine Learning

Gummi Saugnapf Glas Kreisschneider

von Sourcingmap

Geben Sie die erste Bewertung für diesen Artikel ab

Preis: EUR 8,98 **GRATIS Lieferung innerhalb Deutschlands.**

Alle Preisangaben inkl. USt

Nur noch 12 auf Lager

Machine Learning

Machine Learning

"Show my photos from Utah last August"

"What movies are playing today?"

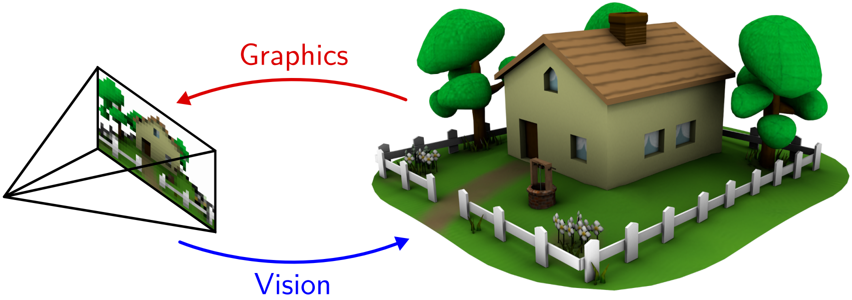
"Get my call history"

Machine Learning

Computer Vision

Computer Vision

2D Image **3D Scene**



Pixel Matrix

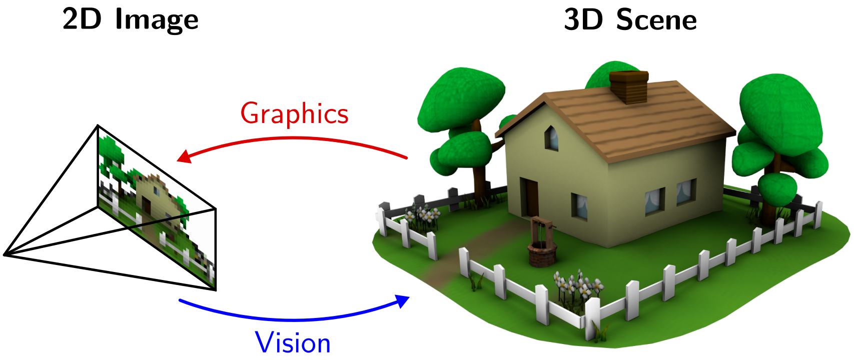
217	191	252	255	239
102	80	200	146	138
159	94	91	121	138
179	106	136	85	41
115	129	83	112	67
94	114	105	111	89

Objects Material

Shape/Geometry Motion

Semantics 3D Pose

Computer Vision



Computer Vision is an ill-posed inverse problem:

- Many 3D scenes yield the same 2D image
- Additional constraints (knowledge about world) required

Why is Computer Vision hard?

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

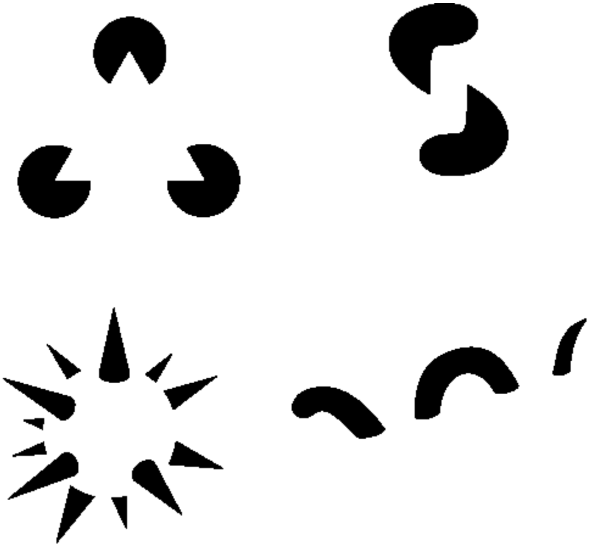
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Why is Visual Perception hard?

Why is Visual Perception hard?

Why is Visual Perception hard?

Why is Visual Perception hard?



Why is Visual Perception hard?

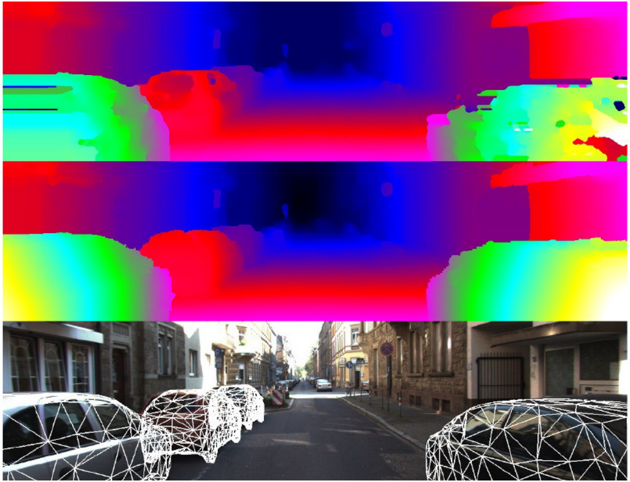
Source: http://www.homeworkshop.com/

Why is Visual Perception hard?

Why is Visual Perception hard?

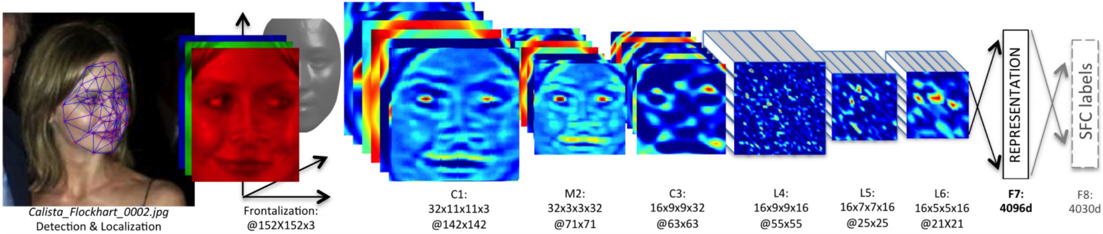
Some Examples

Depth Estimation



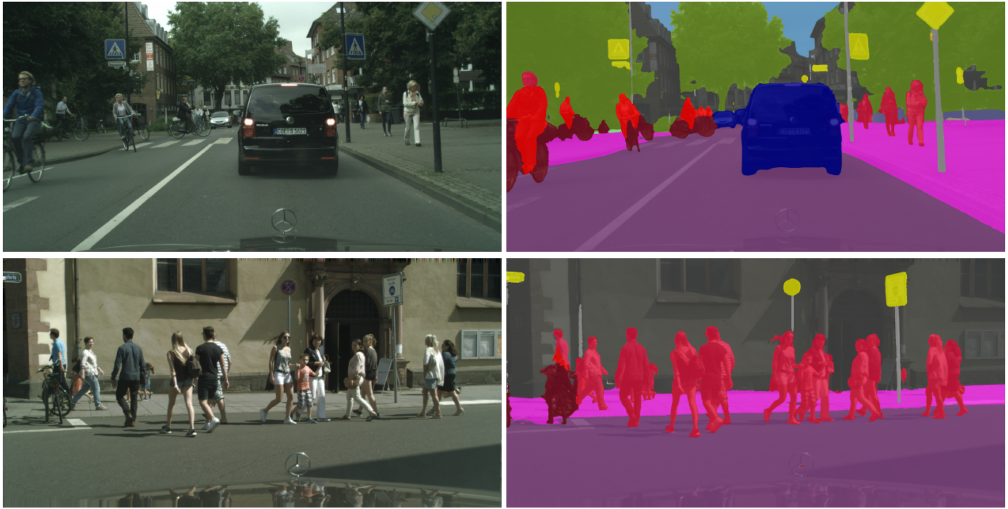
[Güney *et al.*, CVPR 2015]

Face Recognition



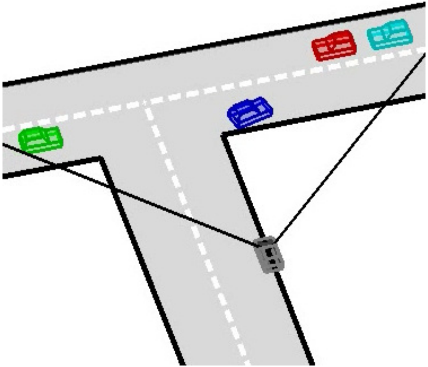
[Taigman *et al.*, CVPR 2014]

Semantic Segmentation

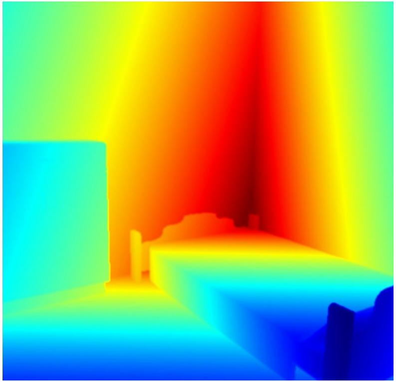


[Kundu *et al.*, CVPR 2016]

Scene Understanding



[Geiger *et al.*, PAMI 2014]



[Geiger & Wang, GCPR 2015]

Image Captioning

"little girl is eating piece of cake."

"baseball player is throwing ball in game."

"woman is holding bunch of bananas."

"black cat is sitting on top of suitcase."

"a young boy is holding a baseball bat."

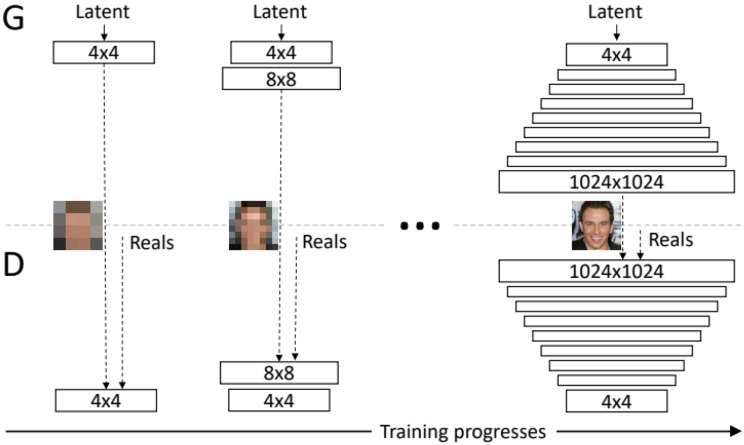
"a cat is sitting on a couch with a remote control."

"a woman holding a teddy bear in front of a mirror."

"a horse is standing in the middle of a road."

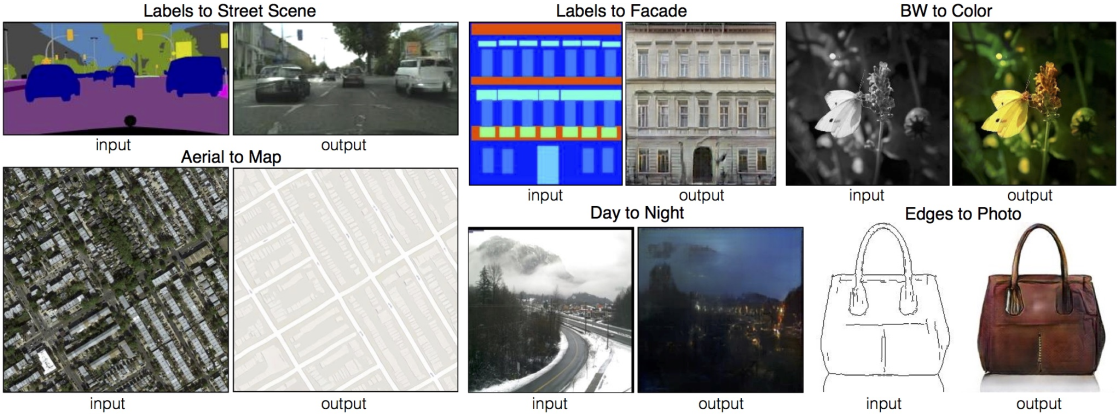
[Karpathy *et al.*, PAMI 2017]

Generative Models



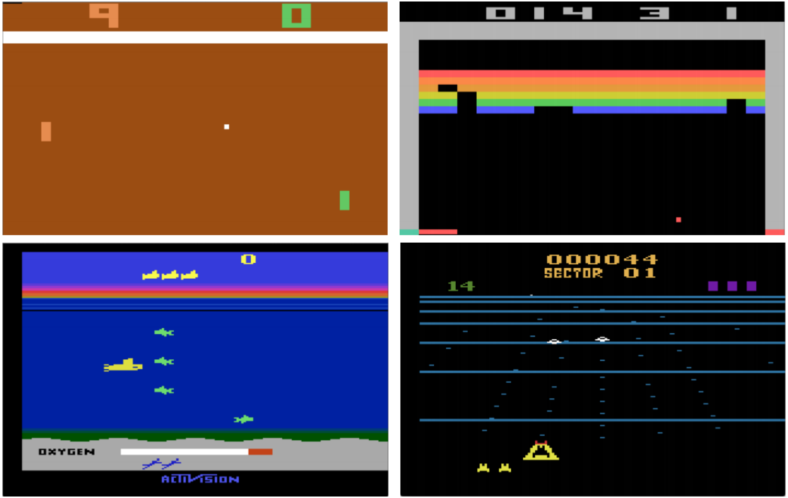
[Karras *et al.*, ICLR 2018]

Image-to-image Translation



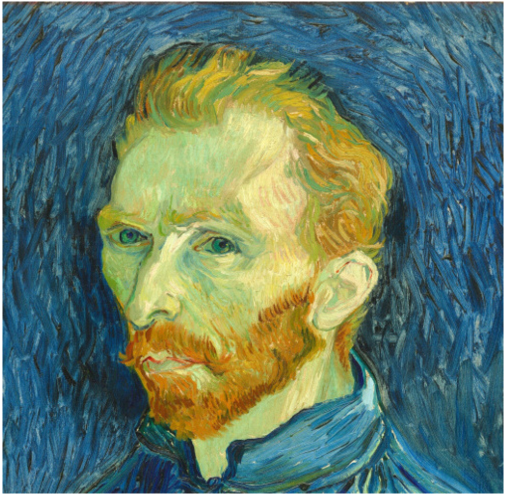
[Isola *et al.*, CVPR 2017]

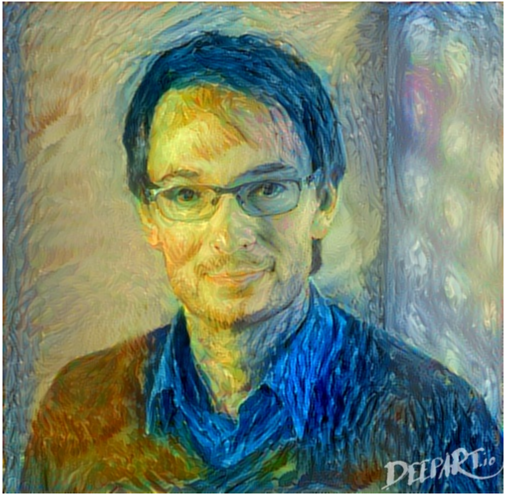
Sensori-Motor Control



[Mnih *et al*., Nature 2015]

Style Transfer

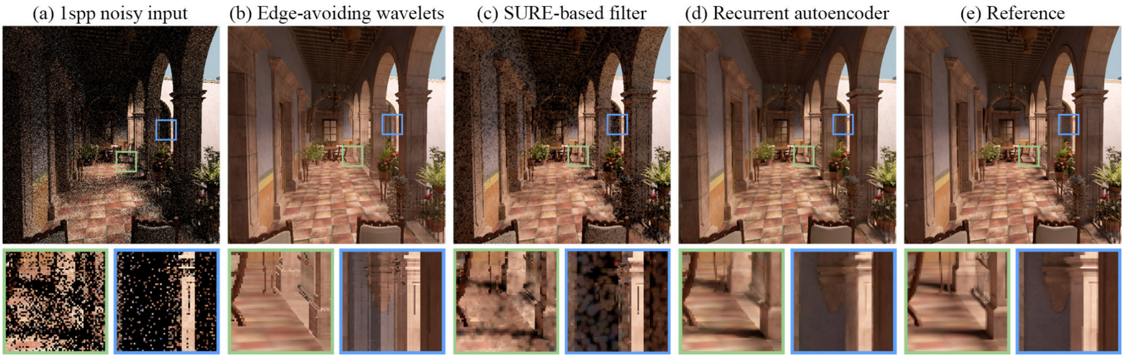




[Gatys *et al.*, CVPR 2016]

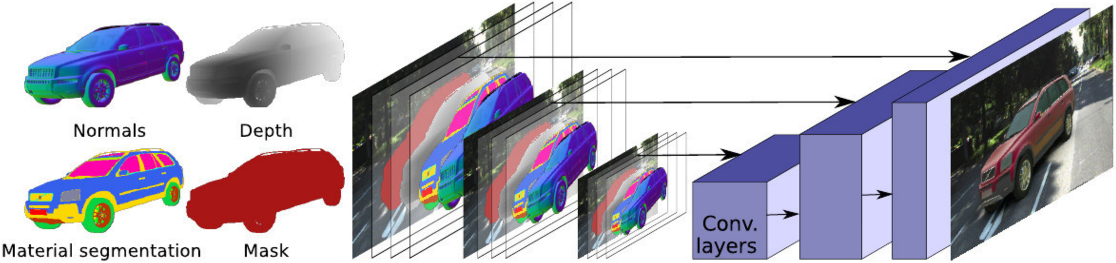
Computer Graphics

Denoising



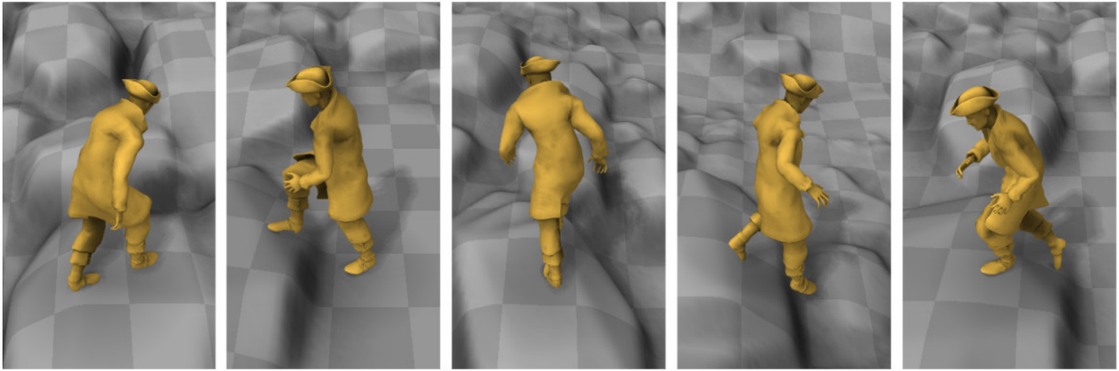
[Chaitanya *et al.*, SIGGRAPH 2017]

Rendering



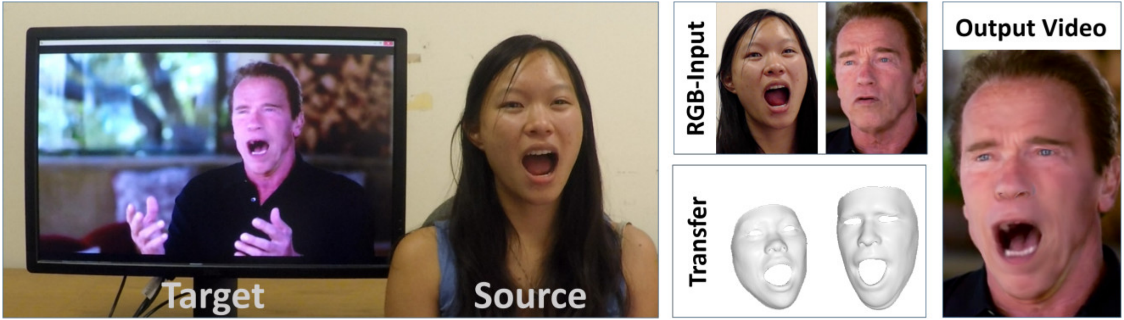
[Alhaija *et al*., 2018]

Animation



[Holden *et al.*, SIGGRAPH 2017]

Reenactment



[Thies *et al*., CVPR 2016]

Machine Learning
Part I

Machine Learning

Problems:

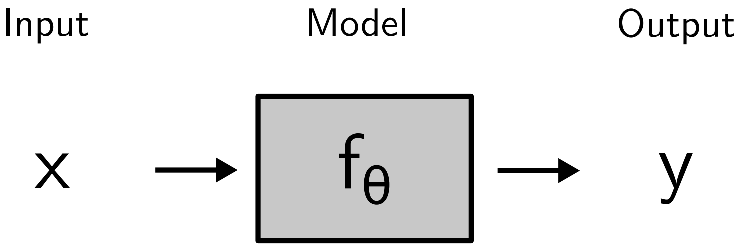
- Supervised learning
  - Classification
  - Regression
  - Structured prediction
- Unsupervised learning
  - Clustering
  - Density Estimation
  - Dimensionality reduction
- Reinforcement learning
- ...

Methods:

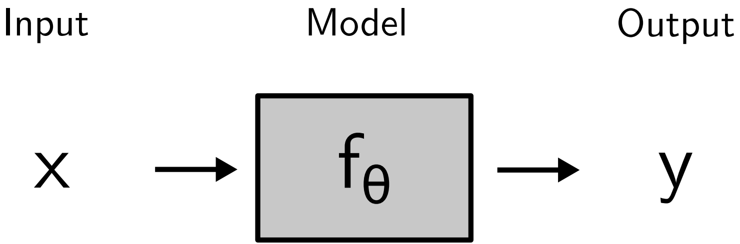
- Sampling
- Gaussian Processes
- Boosting
- Decision Trees
- Support Vector Machines
- Graphical Models
- Deep Neural Networks
- ...

Supervised Learning

Supervised Learning

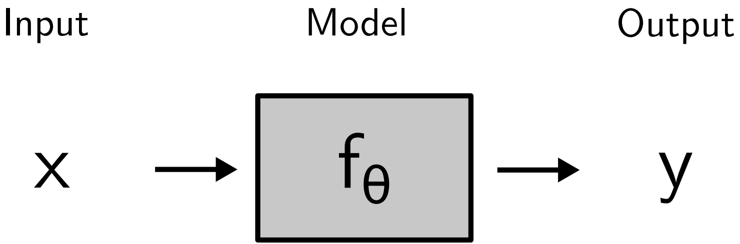


Supervised Learning



- **Learning:** Estimate parameters θ from training data $\{(x_i, y_i)\}_{i=1}^{N}$

Supervised Learning



- **Learning:** Estimate parameters θ from training data $\{(x_i, y_i)\}_{i=1}^N$
- **Inference:** Make novel predictions: $y = f_\theta(x)$

Classification

Input | Model | Output

→ f_θ → "Beach"

- **Mapping:** $f_\theta : \mathbb{R}^{W \times H} \rightarrow$ {"Beach", "No Beach"}

Regression

Input

Model

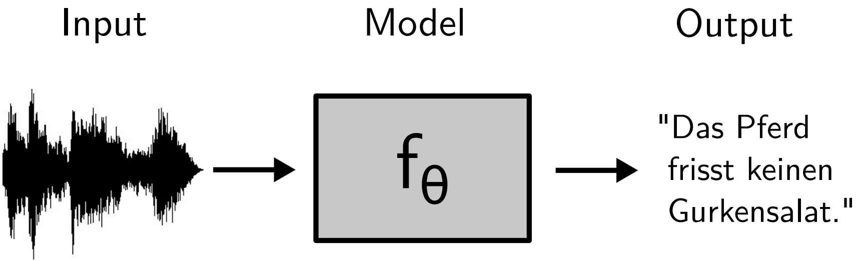
f_θ

Output

143,52 €

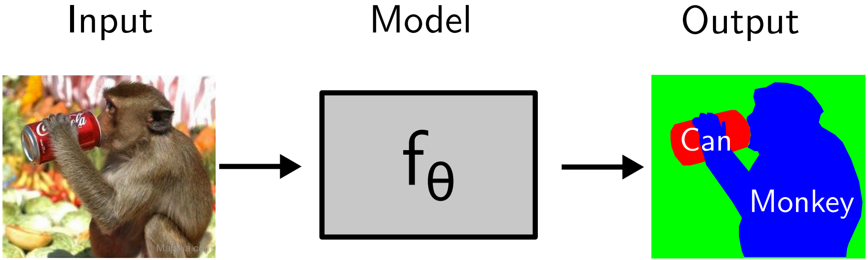
- **Mapping:** $f_\theta : \mathbb{R}^N \to \mathbb{R}$

Structured Prediction



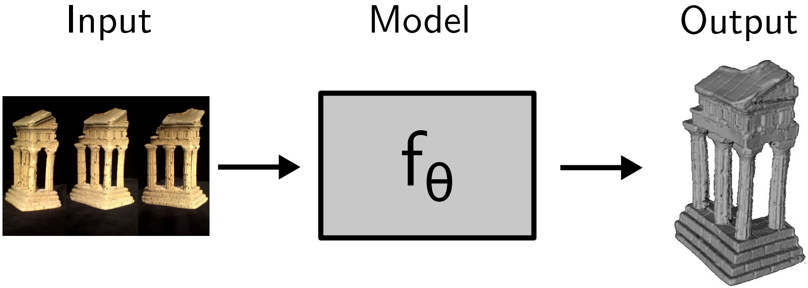
- **Mapping:** $f_\theta : \mathbb{R}^N \rightarrow \{1, \ldots, L\}^M$

Structured Prediction



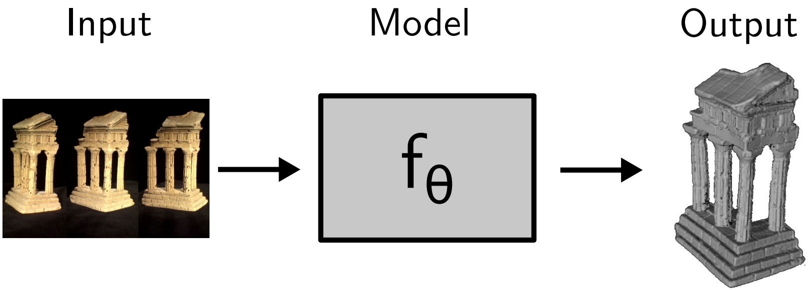
- **Mapping:** $f_\theta : \mathbb{R}^{W \times H} \rightarrow \{1, \ldots, L\}^{W \times H}$

Structured Prediction



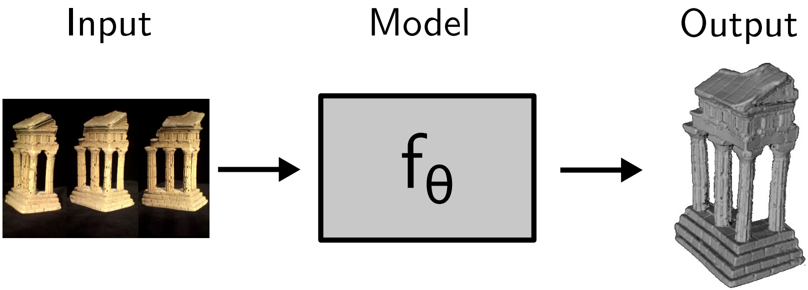
- **Mapping:** $f_\theta : \mathbb{R}^{W \times H \times N} \rightarrow \{0,1\}^{M^3}$

Structured Prediction



- **Mapping:** $f_\theta : \mathbb{R}^{W \times H \times N} \rightarrow \{0,1\}^{M^3}$
- **Suppose:** 32^3 voxels, binary variable per voxel (occupied/free)

Structured Prediction



- **Mapping:** $f_\theta : \mathbb{R}^{W \times H \times N} \rightarrow \{0,1\}^{M^3}$
- **Suppose:** 32^3 voxels, binary variable per voxel (occupied/free)
- **Question:** How many different reconstructions? $2^{32^3} = 2^{32768}$
- **Comparison:** Number of atoms in the universe? $\sim 2^{273}$

Parametric Methods

Parametric Methods

- Define parametric model of the observed process / mapping
- Fit parameters of the model to data
- Make predictions for new inputs using the parameterized model

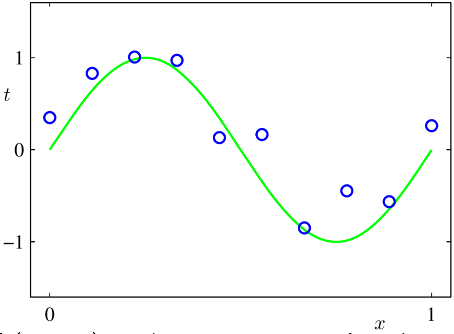
Polynomial Curve Fitting

Polynomial Curve Fitting

Let $\mathcal{X}$ denote a dataset of size N and let $(x_n, t_n) \in \mathcal{X}$ denote its elements.

Here (x_n, t_n) is an input/output pair and we assume $x_n, t_n \in \mathbb{R}$ for simplicity.

Goal: Predict t for a previously unseen input x.



Example with true model (green) and 10 noisy samples drawn from this model (blue).

Polynomial Curve Fitting

Let us choose a polynomial of order M to model the dataset $\mathcal{X}$:

$$y(x, \mathbf{w}) = \sum_{j=0}^{M} w_j x^j$$

Tasks:

- **Training:** Estimate $\mathbf{w}$ from $\mathcal{X}$
- **Inference:** Predict t for novel x given estimated $\mathbf{w}$

Polynomial Curve Fitting

Let us choose a polynomial of order M to model the dataset $\mathcal{X}$:

$$y(x, \mathbf{w}) = \sum_{j=0}^{M} w_j x^j$$

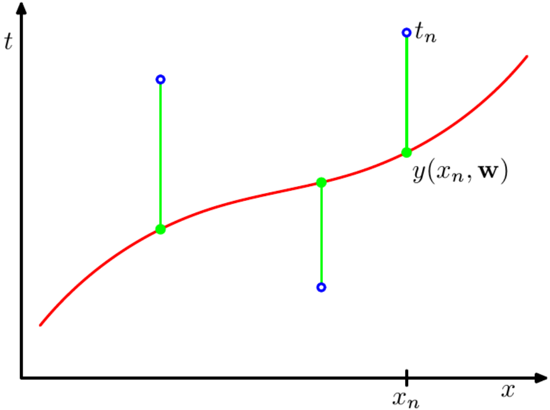
How can we estimate $\mathbf{w}$ from $\mathcal{X}$?

- Need to define a loss function / error function to be minimized, *e.g.*:

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} (y(x_n, \mathbf{w}) - t_n)^2$$

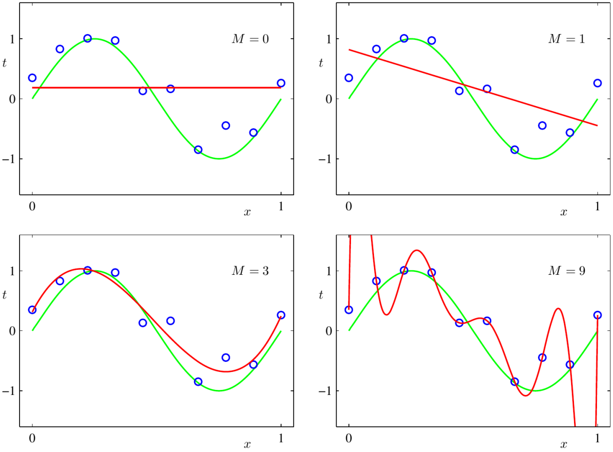
- How can this function be minimized wrt. $\mathbf{w}$? Why is it easy? How many solutions?

Polynomial Curve Fitting



The error function measures the displacement along the t dimension between the data points (blue) and the current model specified by the parameters $\mathbf{w}$ (red).

Polynomial Curve Fitting

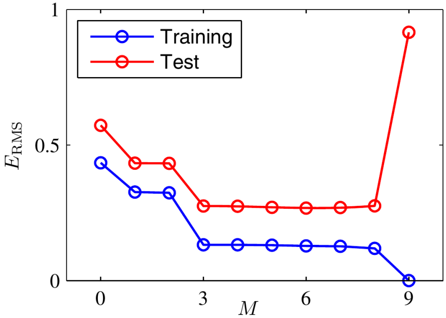


Plots of polynomials of various orders M (red) fitted to the dataset (blue). Underfitting ($M = 0$) and overfitting ($M = 9$) is observed. This is a model selection problem.

Polynomial Curve Fitting

How can we measure generalization performance (overfitting or underfitting)?

- Consider performance on held-out test set from same distribution



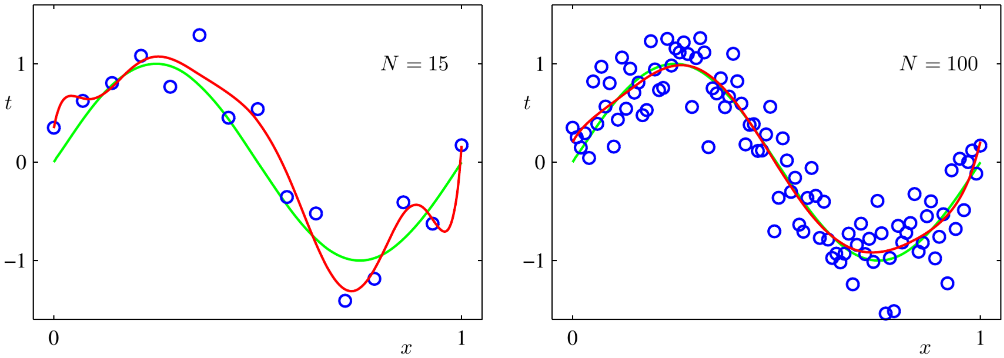
Root Mean Square (RMS) error evaluated on the training set (blue) and test set (red) for various M. Why do we obtain $E_{RMS} = 0$ for $M = 9$ on the training set?

Polynomial Curve Fitting

	$M=0$	$M=1$	$M=6$	$M=9$
$w_0^\star$	0.19	0.82	0.31	0.35
$w_1^\star$		-1.27	7.99	232.37
$w_2^\star$			-25.43	-5321.83
$w_3^\star$			17.37	48568.31
$w_4^\star$				-231639.30
$w_5^\star$				640042.26
$w_6^\star$				-1061800.52
$w_7^\star$				1042400.18
$w_8^\star$				-557682.99
$w_9^\star$				125201.43

Polynomial coefficients $\mathbf{w}$ for polynomials of various order fitted to the dataset.

Polynomial Curve Fitting



Polynomial fits (red) obtained by minimizing the error function using polynomial order $M = 9$ for a dataset of size 15 (left) and 100 (right). What is the effect of more data?

Polynomial Curve Fitting

Model:

$$y(x, \mathbf{w}) = \sum_{j=0}^{M} w_j x^j$$

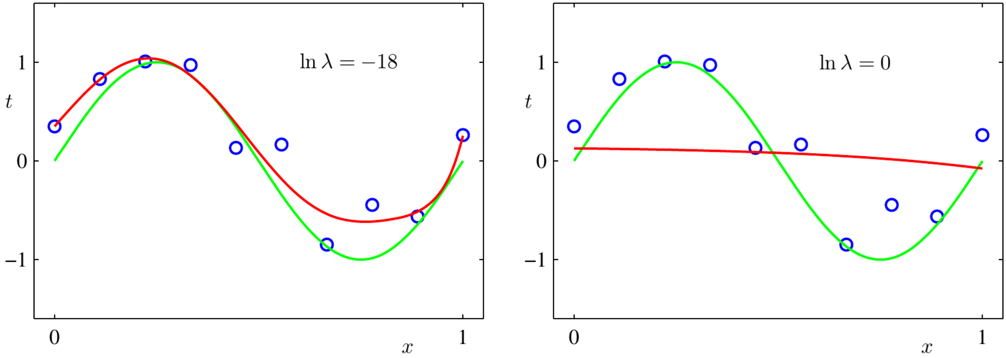
- Loss function:

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} (y(x_n, \mathbf{w}) - t_n)^2$$

- How can we avoid overfitting without reducing K?
- Regularizing the loss function (quadratic case = ridge regression):

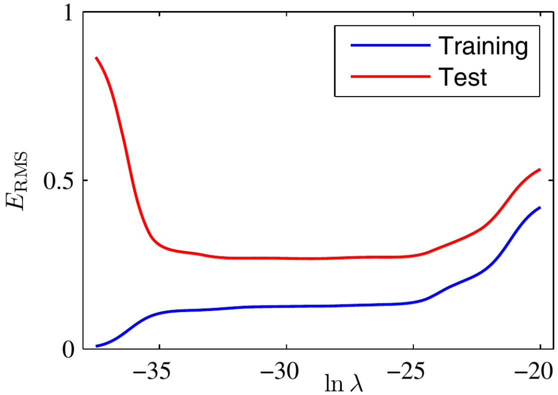
$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} (y(x_n, \mathbf{w}) - t_n)^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

Polynomial Curve Fitting



Regularized polynomial fits (red) with weak (left) and strong (right) regularization.

Polynomial Curve Fitting



Root Mean Square (RMS) error evaluated on the training set (blue) and test set (red) for various regularization values λ. Which λ value would you pick?

Non-Parametric Methods

Parametric vs. Non-Parametric Methods

- Parametric Models
  - Select model with relatively small number of parameters
  - Fit model parameters to data (=training phase)
  - Generalize well if model is correctly specified
  - Do not require the data at inference time
  - Problematic if model is misspecified (*e.g.*, Gaussian for multi-modal distributions)

- Non-Parametric Models
  - Do not require the definition of a parametric model
  - Instead, directly base their predictions on the dataset (often not even training phase)
  - Scale with the size of the dataset
  - Require more storage, often slower
  - However, often worse in terms of generalization (less assumptions specified)

K-Nearest Neighbor Classification

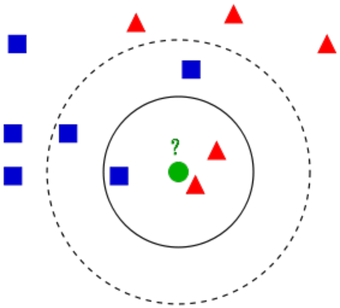
K-NN Classification

K-Nearest Neighbor Classification

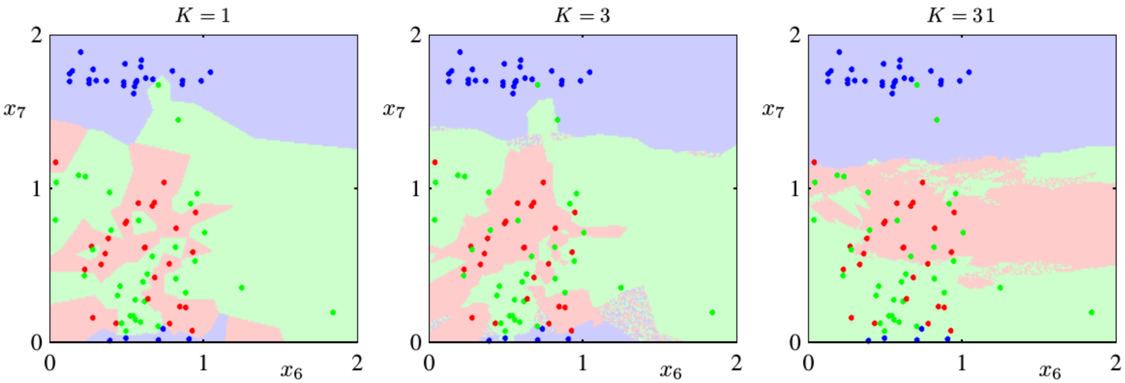
- Assign query point to the class most common among its K nearest neighbors

Special case: Nearest Neighbor Classification

- Assign query point to class of nearest neighbor



K-NN Classification



200 points from the oil dataset (Bishop and James, 1993). This plot shows 2 of 12 feature dimensions for 3 different classes varying K. What do you observe?

Questions?