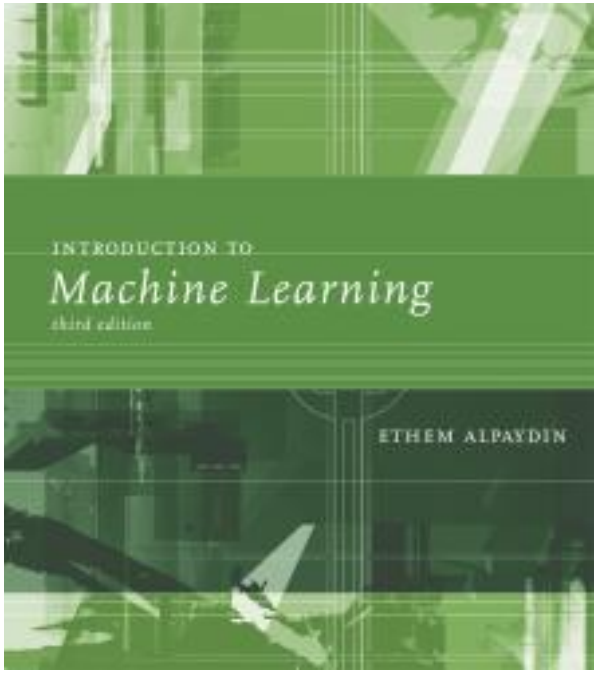
Machine Learning

Artificial Intelligence: A Modern Approach

- Chapter 18: Machine Learning

Optional Textbook: Introduction To Machine Learning

- Used it before



Textbook Author: Ethem Alpaydin

Ethem ALPAYDIN

Türkçe

Professor in the
Department of Computer Engineering
Bogazici University

Member of The Science Academy, Turkey

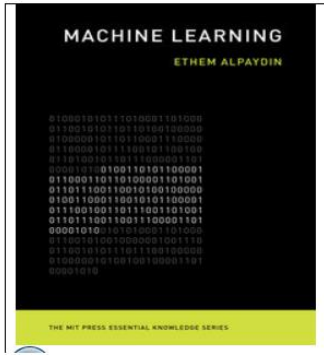
Google Scholar, Semantic Scholar, DBLP, MS Academic Search, ACM Digital Library, Academia.edu, ResearchGate, LinkedIn.

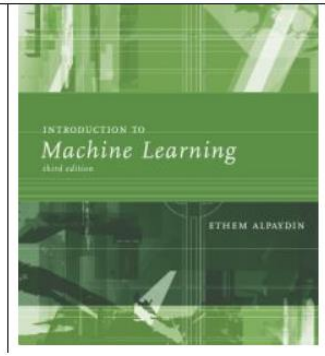
Publications, Talks, Other Writing, Courses, Short Biography, Details

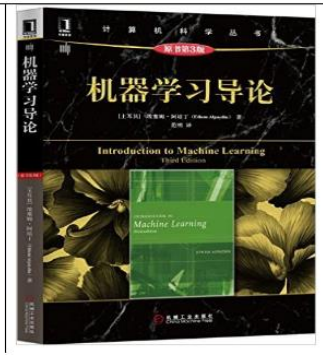
Ethem Alpaydin

+ Follow

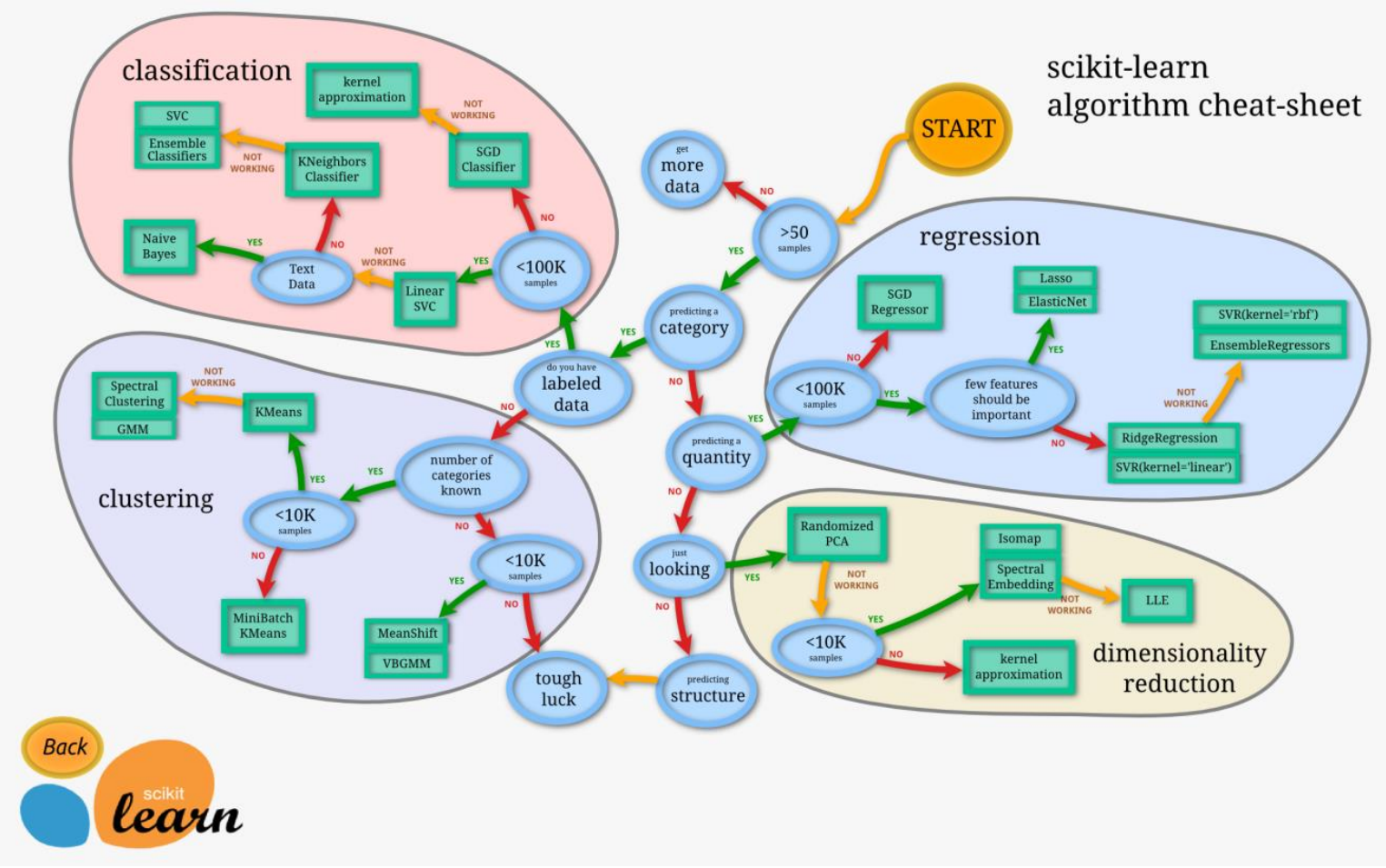
Ethem ALPAYDIN is Professor in the Department of Computer Engineering, Bogazici University, Istanbul Turkey and is a member of the Science Academy, Istanbul. He received his PhD from the Ecole Polytechnique Fédérale de Lausanne, Switzerland in 1990 and was a postdoc at the International Computer Science Institute, Berkeley in 1991. He was a Fulbright scholar in 1997. He was a visiting researcher at MIT, USA in 1994, IDIAP, Switzerland in 1998 and TU Delft, The Netherlands in 2014

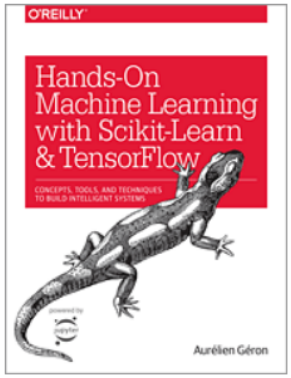






Machine Learning Libraries





Hands-On Machine Learning with Scikit-Learn and TensorFlow

Concepts, Tools, and Techniques to Build Intelligent Systems

By Aurélien Géron

Publisher: O'Reilly Media
Release Date: March 2017
Pages: 576

Graphics in this book are printed in black and white.

Through a series of recent breakthroughs, deep learning has boosted the entire field of machine learning. Now, even programmers who know close to nothing about this technology can use simple, efficient tools to implement programs capable of learning from data. This practical book shows you how.

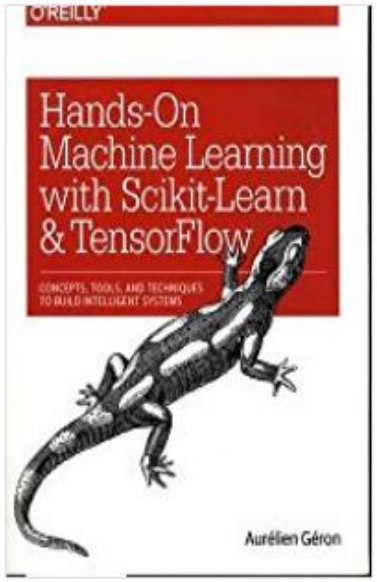
By using concrete examples, minimal theory, and two production-ready Python frameworks—scikit-learn and TensorFlow—author Aurélien Géron helps you gain an intuitive understanding of the concepts and tools for building intelligent systems. You'll learn a range of techniques, starting with simple linear regression and progressing to deep neural networks. With exercises in each chapter to help you apply what you've learned, all you need is programming experience to get started.

- Explore the machine learning landscape, particularly neural nets
- Use scikit-learn to track an example machine-learning project end-to-end
- Explore several training models, including support vector machines, decision trees, random forests, and ensemble methods
- Use the TensorFlow library to build and train neural nets
- Dive into neural net architectures, including convolutional nets, recurrent nets, and deep reinforcement learning
- Learn techniques for training and scaling deep neural nets
- Apply practical code examples without acquiring excessive machine learning theory or algorithm details

OPTIONAL BOOK

Practical Handbook

Optional Book: Author: Aurélien Geron



Aurélien Geron

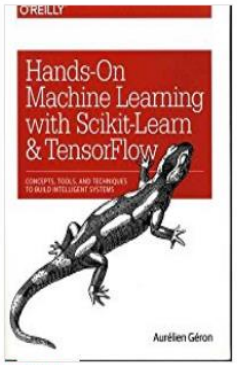
ageron

Machine Learning consultant, former PM of YouTube video classification and founder & CTO of telco operator

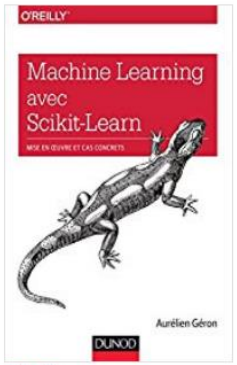
Follow

Aurélien Géron

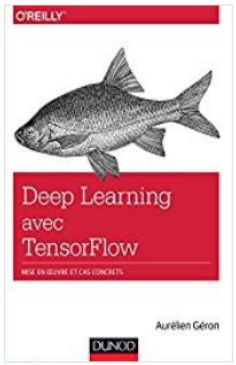
+ Follow



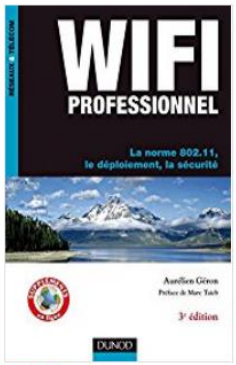
£23.94
Paperback



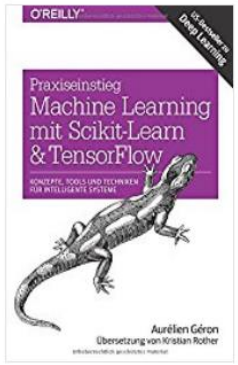
£27.99
Kindle Edition



£31.99
Kindle Edition



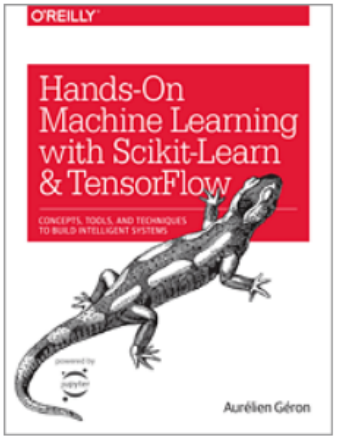
£32.99
Kindle Edition



Paperback

Machine Learning Notebooks

This project aims at teaching you the fundamentals of Machine Learning in python. It contains the example code and solutions to the exercises in my O'Reilly book Hands-on Machine Learning with Scikit-Learn and TensorFlow:

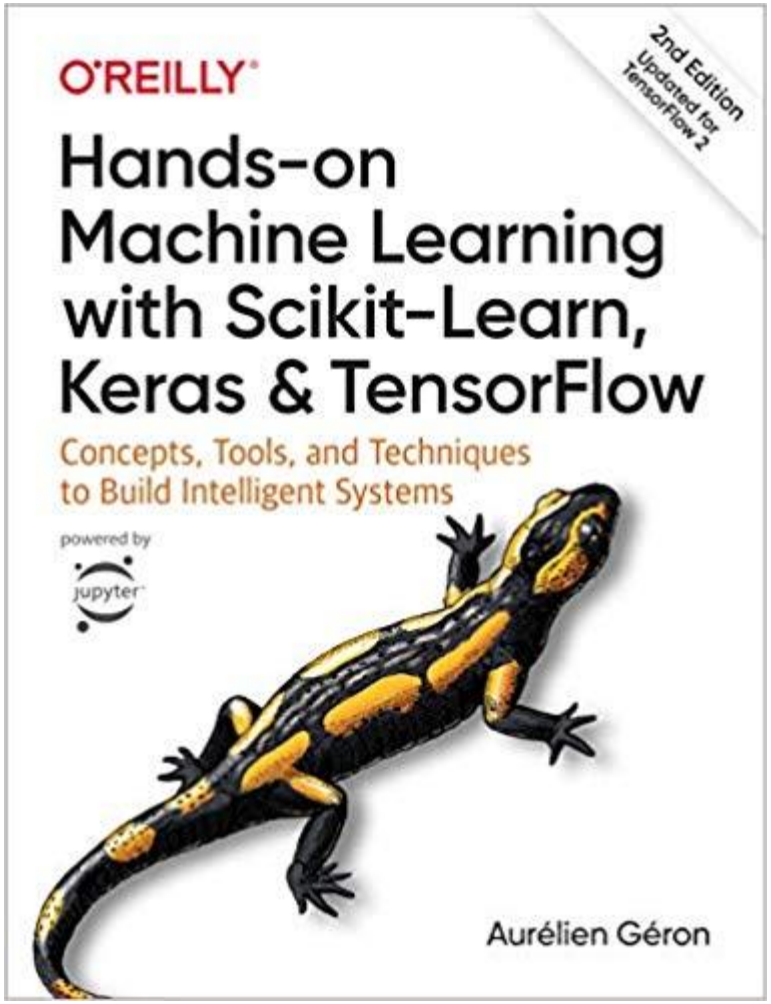


Simply open the Jupyter notebooks you are interested in:

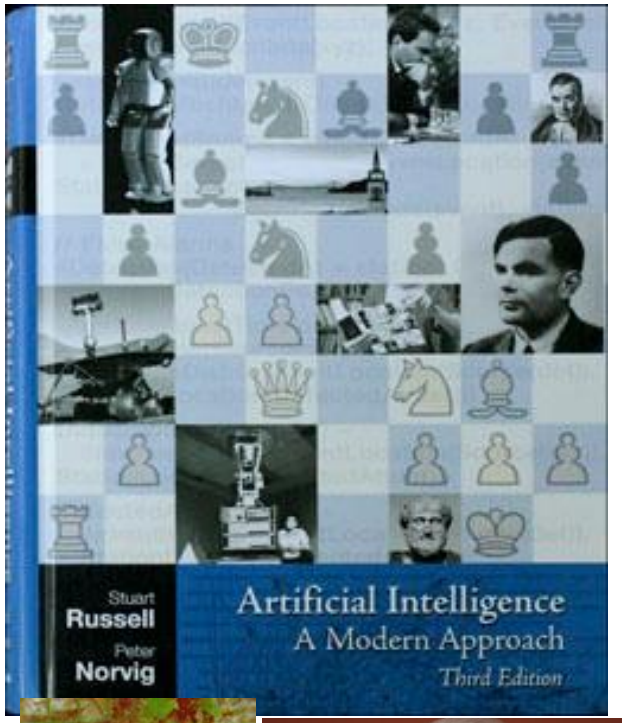
- Using jupyter.org's notebook viewer
  - note: github.com's notebook viewer also works but it is slower and the math formulas are not displayed correctly,
- or by cloning this repository and running Jupyter locally. This option lets you play around with the code. In this case, follow the installation instructions below.

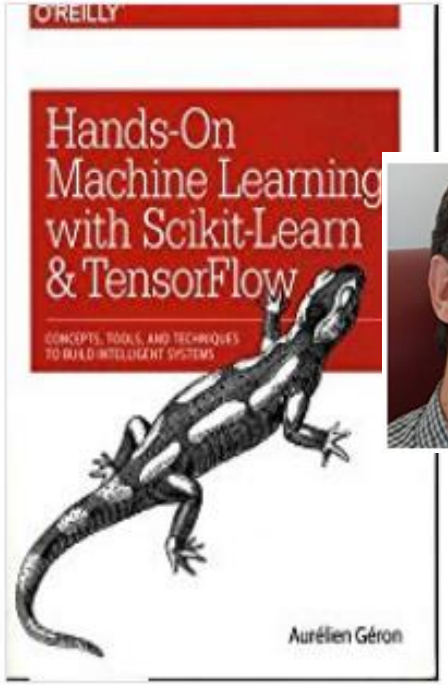
New Second Edition

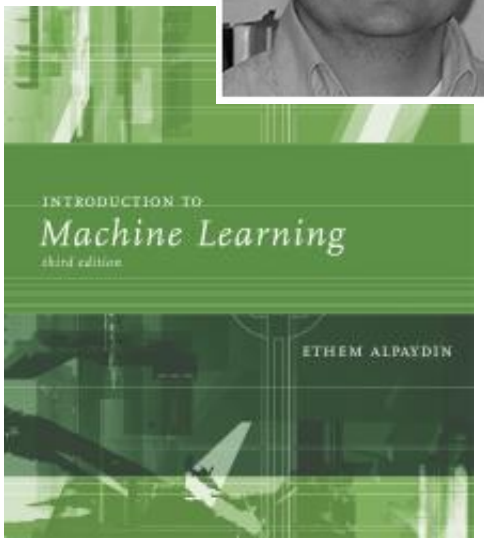
- Tensor Flow 2.0



3 Sources --- 3 Perspectives







Big Data

- Widespread use of personal computers and wireless communication leads to “big data”
- We are both producers and consumers of data
- Data is not random, it has structure, e.g., customer behavior
- We need “big theory” to extract that structure from data for

  (a) Understanding the process

  (b) Making predictions for the future

Why "Learn" ?

- Machine learning is programming computers to optimize a performance criterion using example data or past experience.
- There is no need to "learn" to calculate payroll
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

What We Talk About When We Talk About "Learning"

- Learning general models from a data of particular examples
- Data is cheap and abundant (data warehouses, data marts); knowledge is expensive and scarce.
- Example in retail: Customer transactions to consumer behavior:

  People who bought "Blink" also bought "Outliers" (www.amazon.com)

- Build a model that is *a good and useful approximation* to the data.

Data Mining

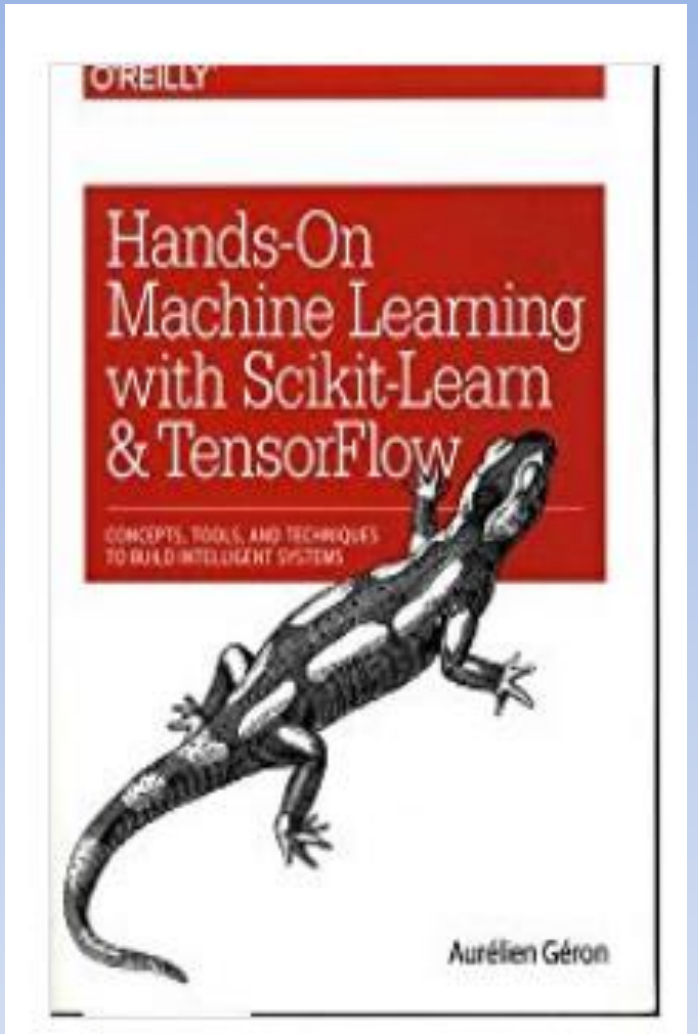
- Retail: Market basket analysis, Customer relationship management (CRM)
- Finance: Credit scoring, fraud detection
- Manufacturing: Control, robotics, troubleshooting
- Medicine: Medical diagnosis
- Telecommunications: Spam filters, intrusion detection
- Bioinformatics: Motifs, alignment
- Web mining: Search engines
- …

What is Machine Learning?

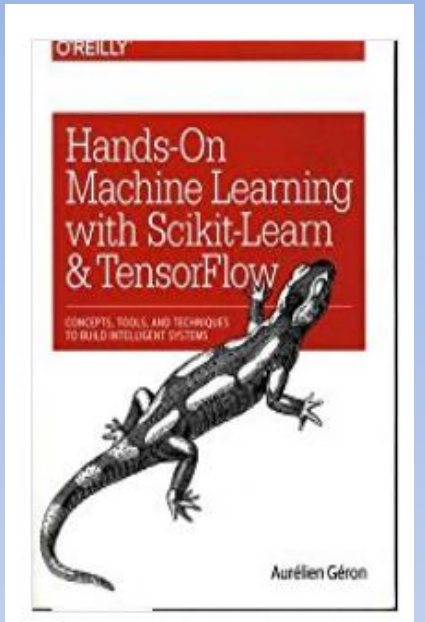
- Optimize a performance criterion using example data or past experience.
- Role of Statistics: Inference from a sample
- Role of Computer science: Efficient algorithms to
  - Solve the optimization problem
  - Representing and evaluating the model for inference

Hands-On
Why Machine Learning?

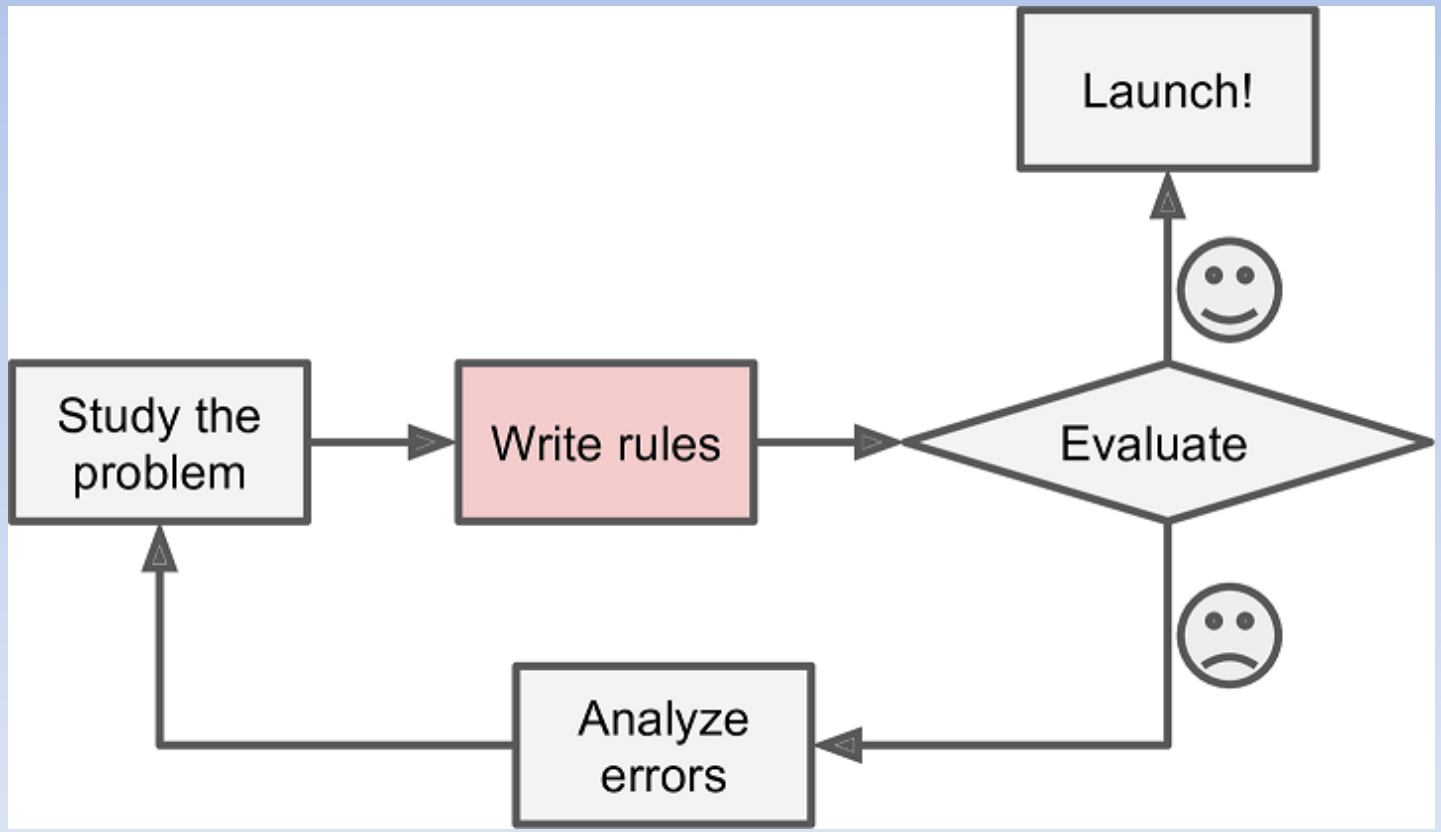
- More Practical Text



Hands-On
Why Machine Learning?

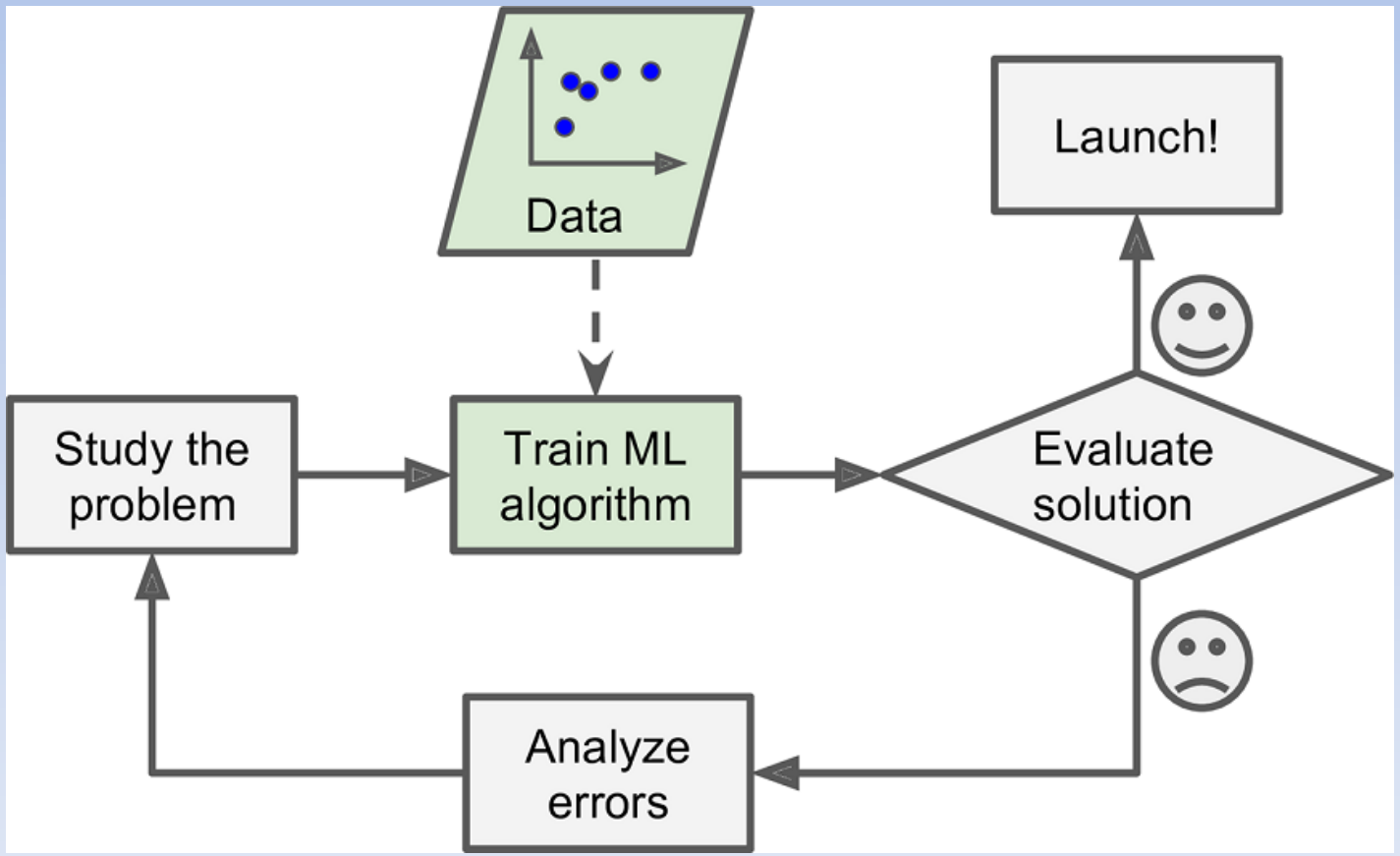


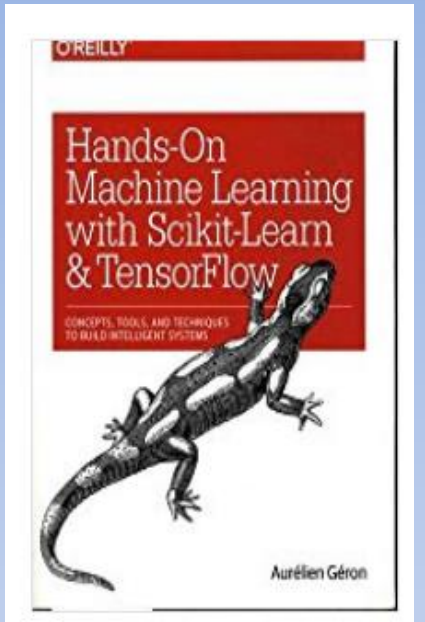
- Traditional Programming



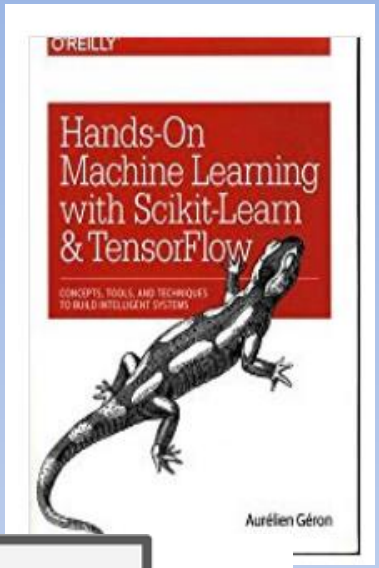
Hands-On
Why Machine Learning?

- Machine Learning Approach

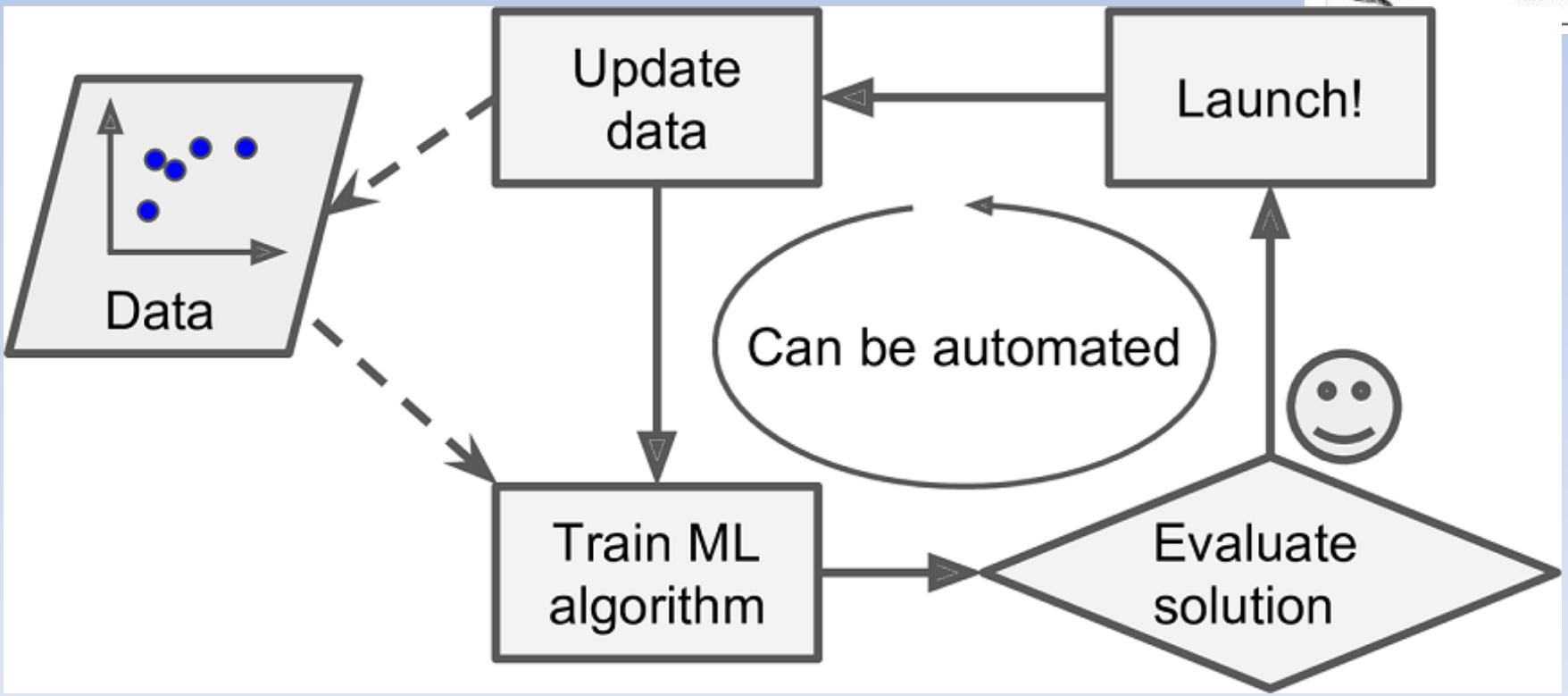




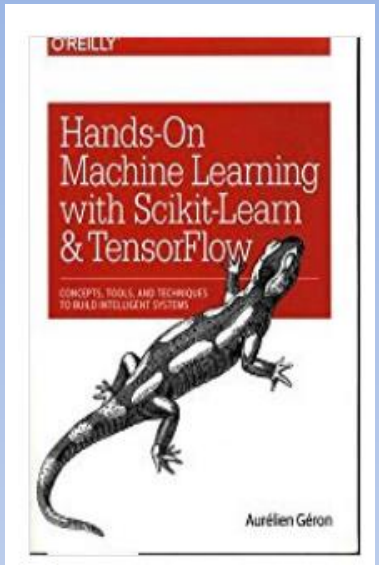
Hands-On
Why Machine Learning?



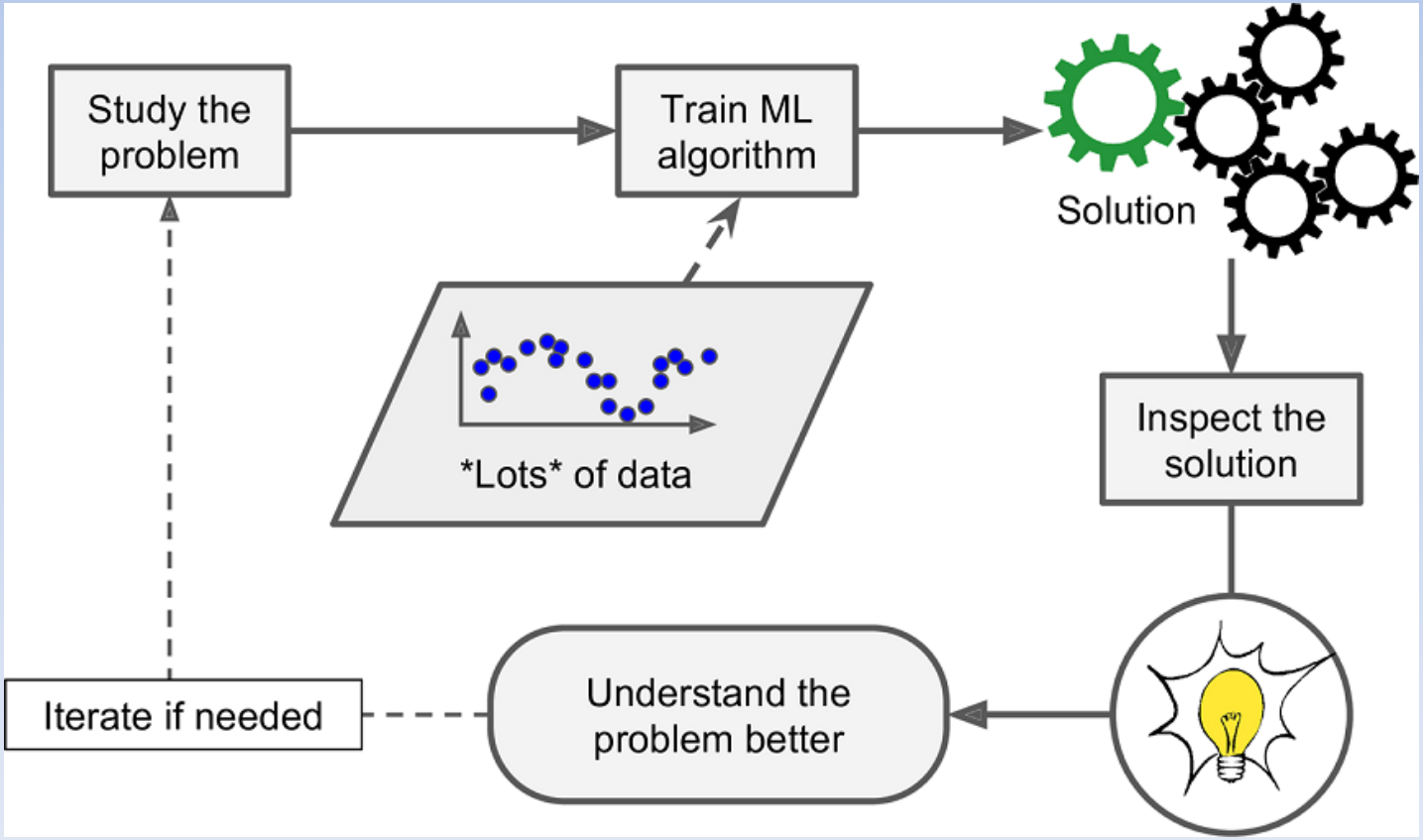
- Machine Learning Can Change Automatically!



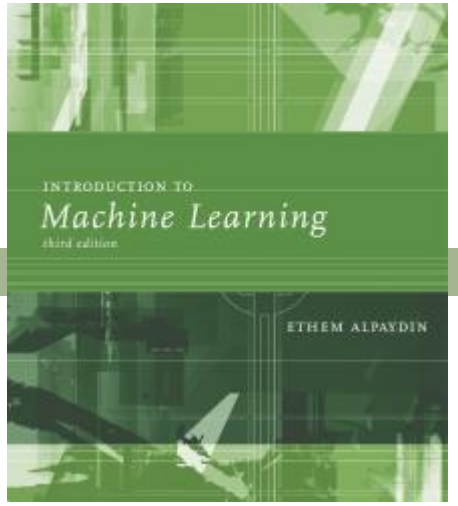
Hands-On
Why Machine Learning?



- Machine Learning Can Aid Understanding.



Applications / Types of Machine Learning



- Association
- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
- Reinforcement Learning

Learning Associations

- Basket analysis:

  $P(Y \mid X)$ probability that somebody who buys X also buys Y where X and Y are products/services.

  Example: $P(\text{chips} \mid \text{beer}) = 0.7$

Market Basket Analysis

- https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce

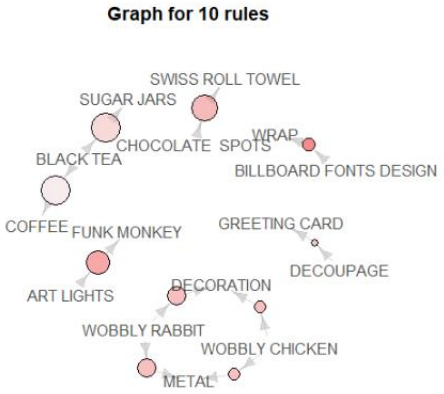
Machine Learning Repository
Center for Machine Learning and Intelligent Systems

Online Retail Data Set

Download: Data Folder, Data Set Description

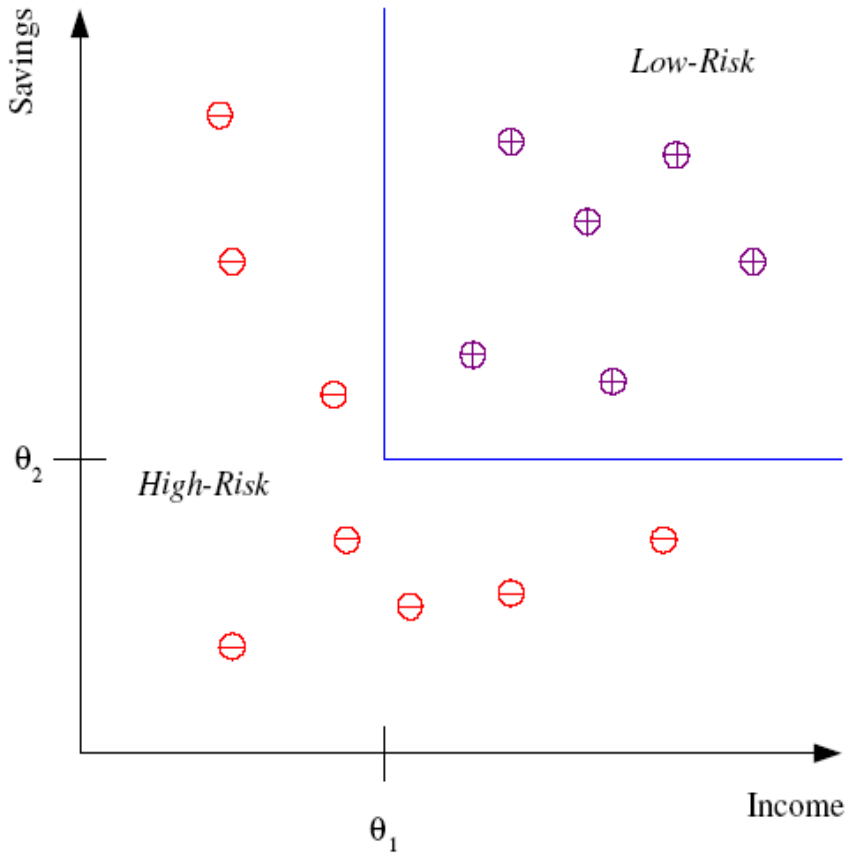
Abstract: This is a transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.

Data Set Characteristics:	Multivariate, Sequential, Time-Series	**Number of Instances:**	541909	**Area:**	Business
Attribute Characteristics:	Integer, Real	**Number of Attributes:**	8	**Date Donated**	2015-11-06
Associated Tasks:	Classification, Clustering	**Missing Values?**	N/A	**Number of Web Hits:**	288593



Supervised Learning: Classification

- Example: Credit scoring
- Differentiating between low-risk and high-risk customers from their *income* and *savings*



Discriminant: IF *income* > θ_1 AND *savings* > θ_2
THEN low-risk ELSE high-risk

Classification: Applications

- Aka Pattern recognition
- Face recognition: Pose, lighting, occlusion (glasses, beard), make-up, hair style
- Character recognition: Different handwriting styles.
- Speech recognition: Temporal dependency.
- Medical diagnosis: From symptoms to illnesses
- **Biometrics:** Recognition/authentication using physical and/or behavioral characteristics: Face, iris, signature, etc
- Outlier/novelty detection:

Face Recognition

Training examples of a person

Test images

ORL dataset,
AT&T Laboratories, Cambridge UK

Biometrics

Unique Identification Authority of India
Government of India

Mera Aadhaar, Meri Pehchaan

AADHAAR

Home | About UIDAI | Legal Framework | Your Aadhaar | Enrolment & Update | Authentication | Media Centre | Resources | Click to search

AADHAAR IS THE WORLD'S LARGEST BIOMETRIC ID PROGRAM

WHAT'S NEW अवैध निवासियों के आधार रद्द होंगे

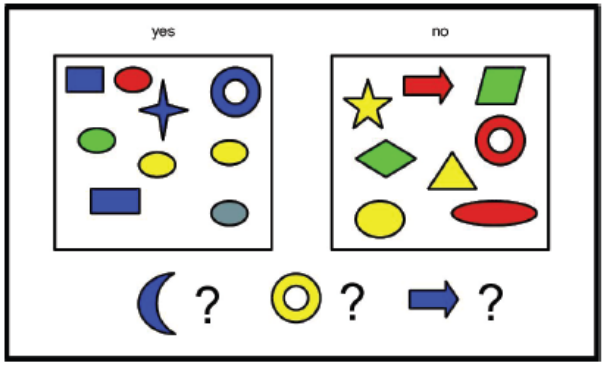
Total Aadhaar Generated 123,14,67,029

About Aadhaar

Aadhaar number is a 12-digit random number issued by the UIDAI ("Authority") to the residents of India after satisfying the verification process laid down by the Authority. Any individual, irrespective of age and gender, who is a resident of India, may voluntarily enrol to obtain Aadhaar number. Person willing to enrol has to provide minimal demographic and biometric information during the enrolment process which is totally free of cost. An individual needs to enrol for Aadhaar only once and after de-duplication **only one Aadhaar** shall be generated, as the **uniqueness** is achieved through the process of demographic and biometric de-duplication.

Demographic information	Name, Date of Birth (verified) or Age (declared), Gender, Address, Mobile Number (optional) and Email ID (optional)
Biometric information	Ten Fingerprints, Two Iris Scans, and Facial Photograph

Classification: R&N CH18



(a)

D features (attributes)

N cases

Color	Shape	Size (cm)	Label
Blue	Square	10	1
Red	Ellipse	2.4	1
Red	Ellipse	20.7	0

(b)

- Learn a function from examples
- f is the target function
- An example is an input-output pair: (x, f(x))

- Problem:
  - Given a hypothesis space *H*
  - Given a training set of examples: $(x_1, f(x_1)), \ldots, (x_n, f(x_n))$
  - Find a hypothesis $h(x)$ such that $h \sim f$

Restaurant Example w/ R&N CH18

- Wait for a table????
- Goal Predicate: WillWait
- Attributes:
  - **Alternate:** whether there is a suitable alternative restaurant nearby.
  - **Bar:** whether the restaurant has a comfortable bar area to wait in.
  - **Fri/Sat:** true on Fridays and Saturdays.
  - **Hungry:** whether we are hungry.
  - **Patrons:** how many people are in the restaurant (values are None, Some, and Full).
  - **Price:** the restaurant's price range ($, $$, $$$).
  - **Raining:** whether it is raining outside.
  - **Reservation:** whether we made a reservation.
  - **Type:** the kind of restaurant (French, Italian, Thai, or burger).
  - **WaitEstimate:** the wait estimated by the host (0–10 minutes, 10–30, 30–60, or >60).

Restaurant Example

Example	Input Attributes										Goal
	Alt	*Bar*	*Fri*	*Hun*	*Pat*	*Price*	*Rain*	*Res*	*Type*	*Est*	*WillWait*
$\mathbf{x}_1$	*Yes*	*No*	*No*	*Yes*	*Some*	*$$$*	*No*	*Yes*	*French*	*0–10*	$y_1 = Yes$
$\mathbf{x}_2$	*Yes*	*No*	*No*	*Yes*	*Full*	*$*	*No*	*No*	*Thai*	*30–60*	$y_2 = No$
$\mathbf{x}_3$	*No*	*Yes*	*No*	*No*	*Some*	*$*	*No*	*No*	*Burger*	*0–10*	$y_3 = Yes$
$\mathbf{x}_4$	*Yes*	*No*	*Yes*	*Yes*	*Full*	*$*	*Yes*	*No*	*Thai*	*10–30*	$y_4 = Yes$
$\mathbf{x}_5$	*Yes*	*No*	*Yes*	*No*	*Full*	*$$$*	*No*	*Yes*	*French*	*>60*	$y_5 = No$
$\mathbf{x}_6$	*No*	*Yes*	*No*	*Yes*	*Some*	*$$*	*Yes*	*Yes*	*Italian*	*0–10*	$y_6 = Yes$
$\mathbf{x}_7$	*No*	*Yes*	*No*	*No*	*None*	*$*	*Yes*	*No*	*Burger*	*0–10*	$y_7 = No$
$\mathbf{x}_8$	*No*	*No*	*No*	*Yes*	*Some*	*$$*	*Yes*	*Yes*	*Thai*	*0–10*	$y_8 = Yes$
$\mathbf{x}_9$	*No*	*Yes*	*Yes*	*No*	*Full*	*$*	*Yes*	*No*	*Burger*	*>60*	$y_9 = No$
$\mathbf{x}_{10}$	*Yes*	*Yes*	*Yes*	*Yes*	*Full*	*$$$*	*No*	*Yes*	*Italian*	*10–30*	$y_{10} = No$
$\mathbf{x}_{11}$	*No*	*No*	*No*	*No*	*None*	*$*	*No*	*No*	*Thai*	*0–10*	$y_{11} = No$
$\mathbf{x}_{12}$	*Yes*	*Yes*	*Yes*	*Yes*	*Full*	*$*	*No*	*No*	*Burger*	*30–60*	$y_{12} = Yes$

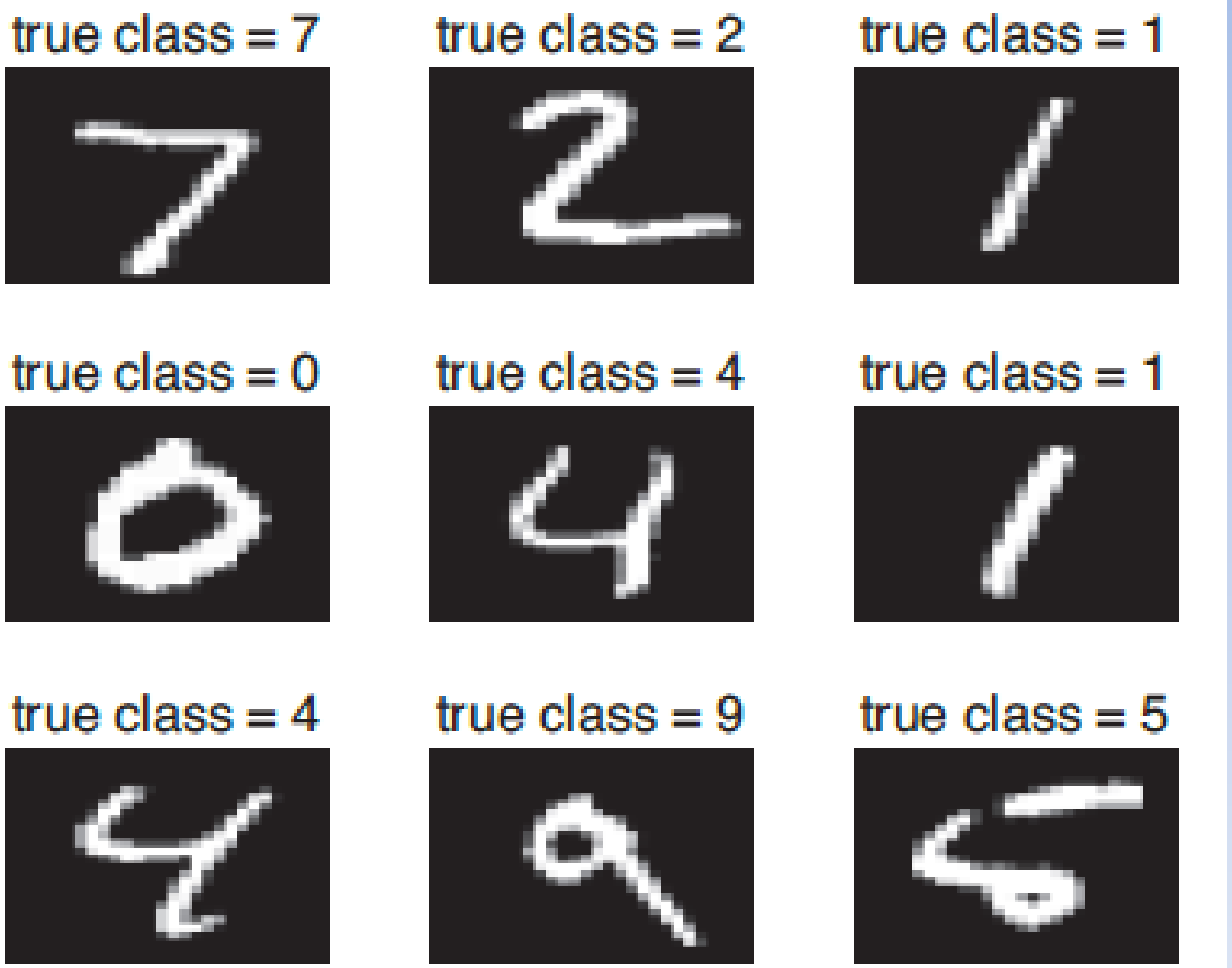
Figure 18.3 Examples for the restaurant domain.

Flowers

(a) (b) (c)

Figure 1.3 Three types of iris flowers: setosa, versicolor and virginica. Source: `http://www.statlab.uni-heidelberg.de/data/iris/` . Used with kind permission of Dennis Kramb and SIGNA.

Image Classification



Regression

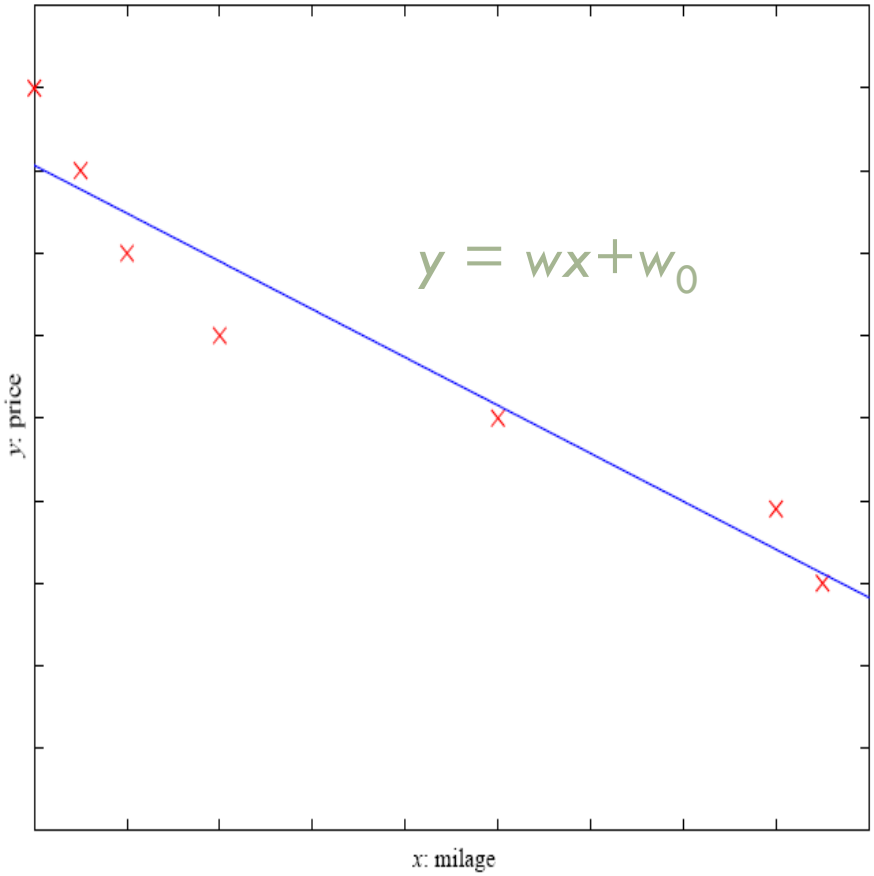
- Example: Price of a used car
- x : car attributes

  y : price

  $$y = g(x \mid \theta)$$

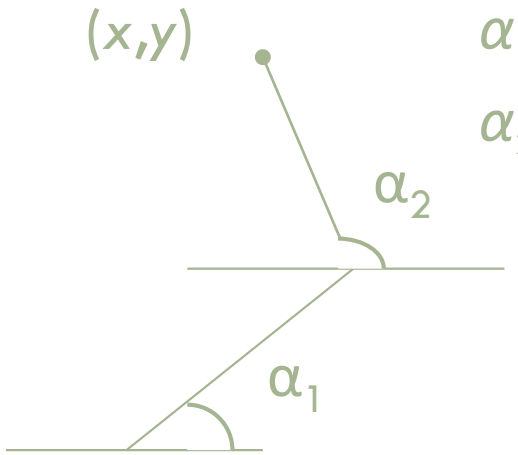
  $g(\ )$ model,

  θ parameters



Regression Applications

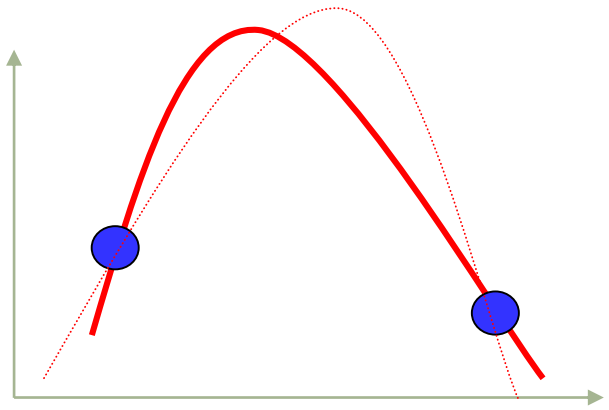
- Navigating a car: Angle of the steering
- Kinematics of a robot arm



$\alpha_1 = g_1(x,y)$

$\alpha_2 = g_2(x,y)$

- Response surface design



Supervised Learning: Uses

- Prediction of future cases: Use the rule to predict the output for future inputs
- Knowledge extraction: The rule is easy to understand
- Compression: The rule is simpler than the data it explains
- Outlier detection: Exceptions that are not covered by the rule, e.g., fraud

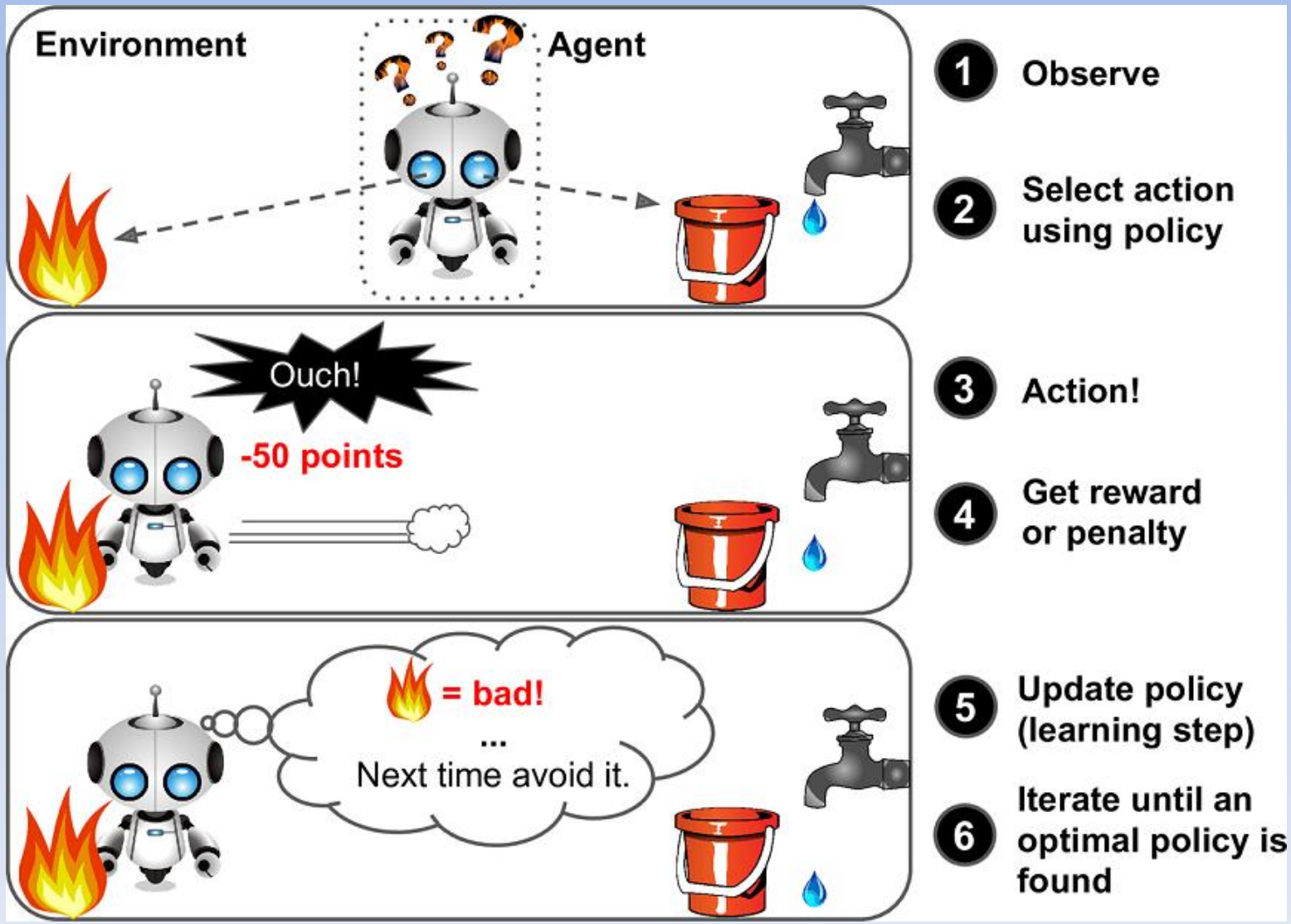
Unsupervised Learning

- Learning “what normally happens”
- No output
- Clustering: Grouping similar instances
- Example applications
  - Customer segmentation in CRM
  - Image compression: Color quantization
  - Bioinformatics: Learning motifs

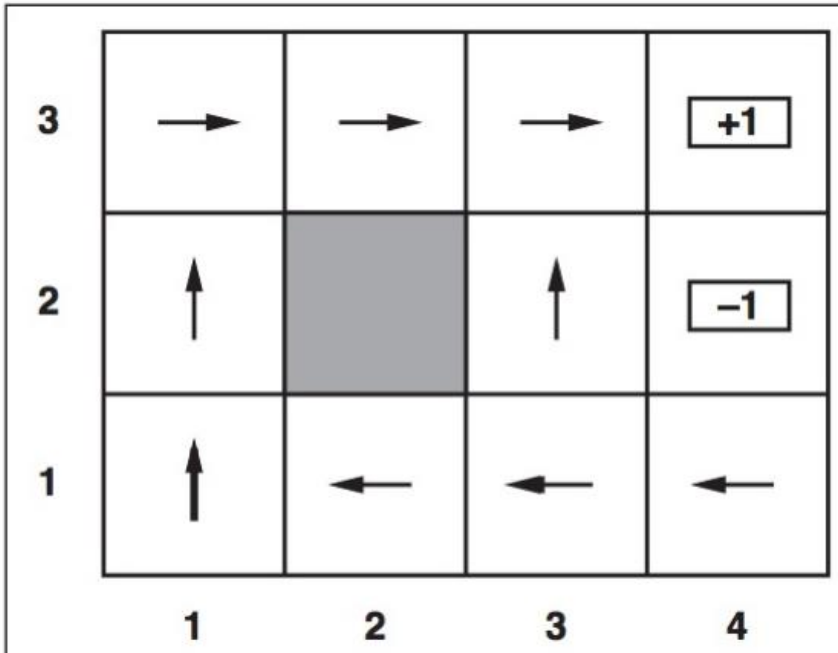
Reinforcement Learning

- Learning a policy: A sequence of outputs
- No supervised output but delayed reward
- Credit assignment problem
- Game playing
- Robot in a maze
- Multiple agents, partial observability, ...

Reinforcement Learning w/ Hands-On



Reinforcement Learning
R&N Chapter 21



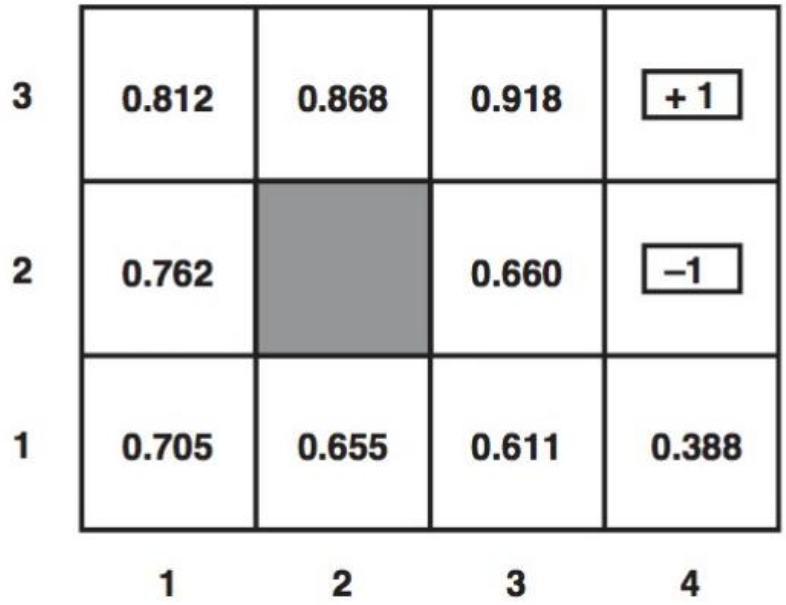
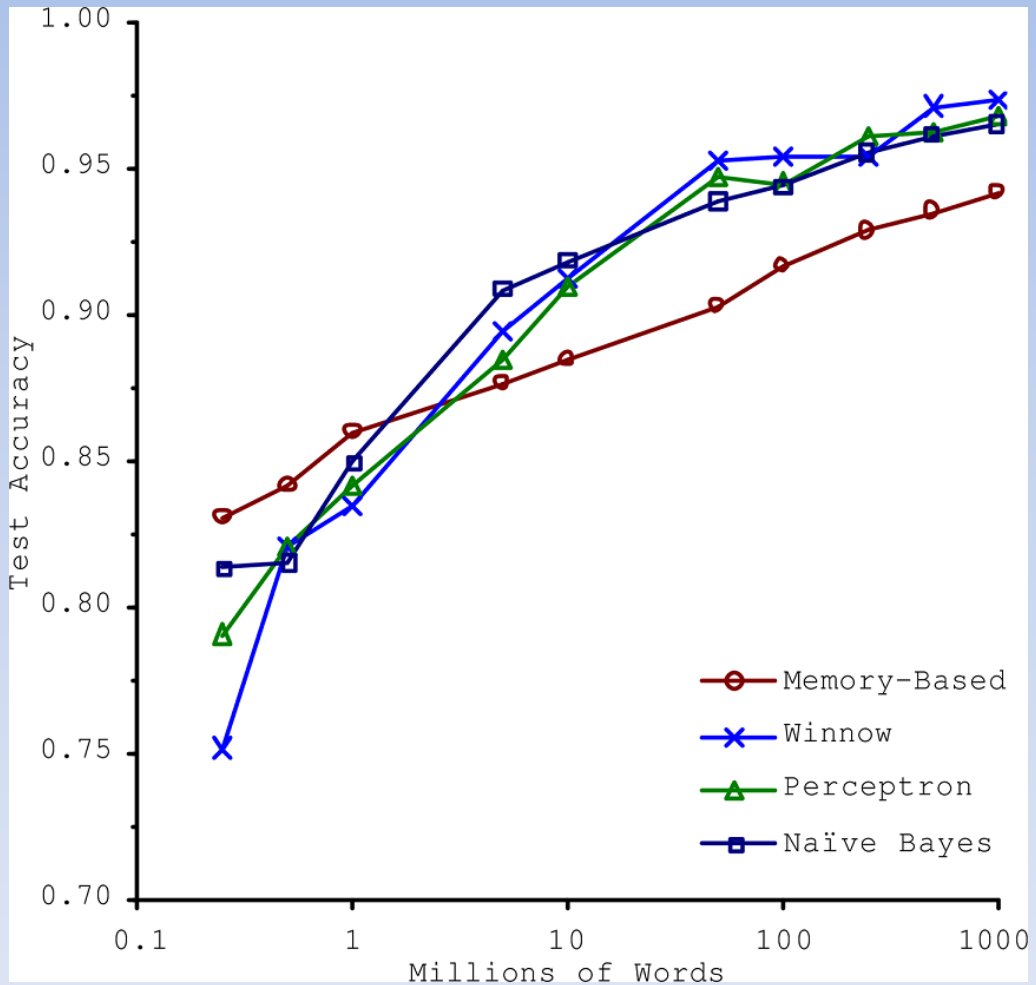


Figure 21.1 (a) A policy π for the 4×3 world; this policy happens to be optimal with rewards of $R(s) = -0.04$ in the nonterminal states and no discounting. (b) The utilities of the states in the 4×3 world, given policy π.

Challenges

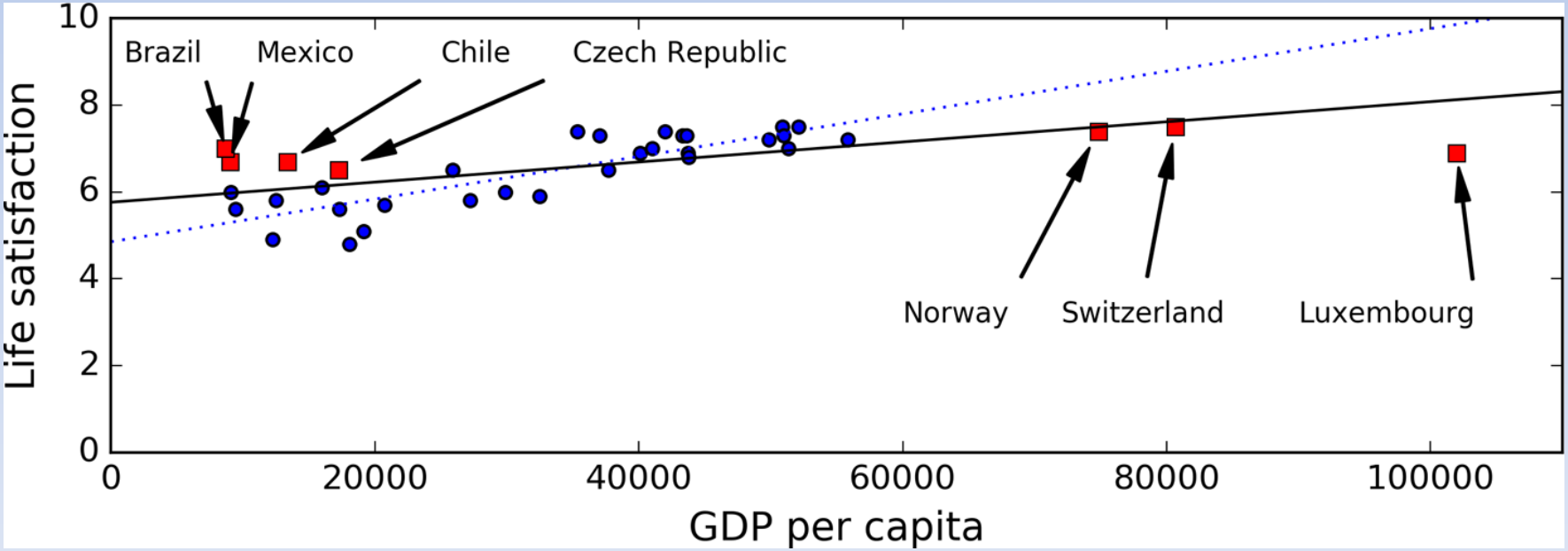
Challenges: Insufficient Quantity of Training Data

- Many Algorithms do well with enough data!

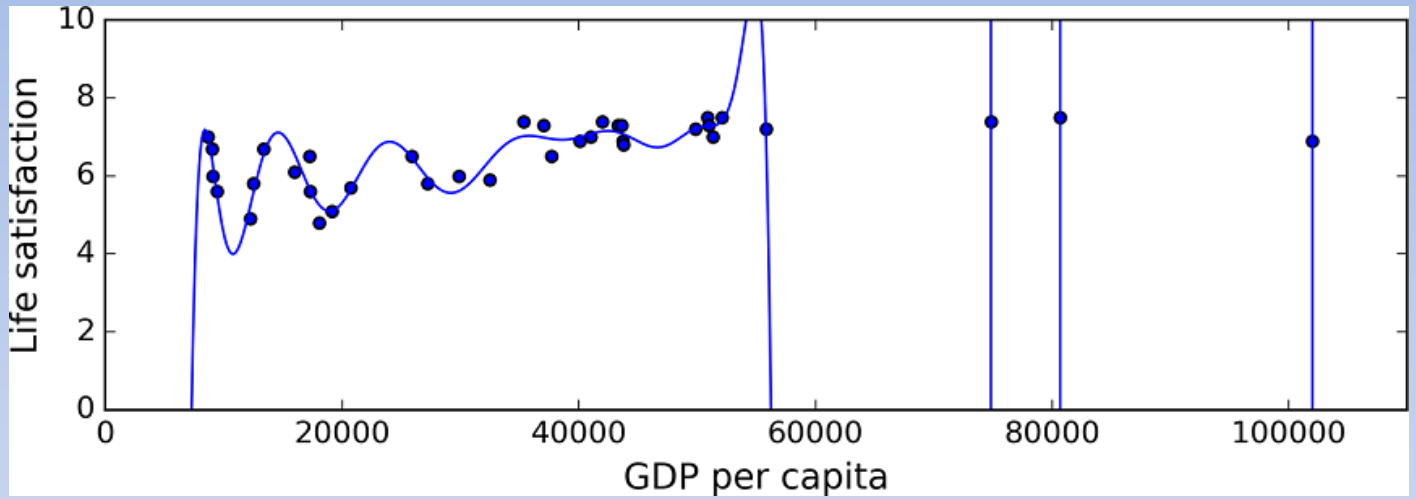


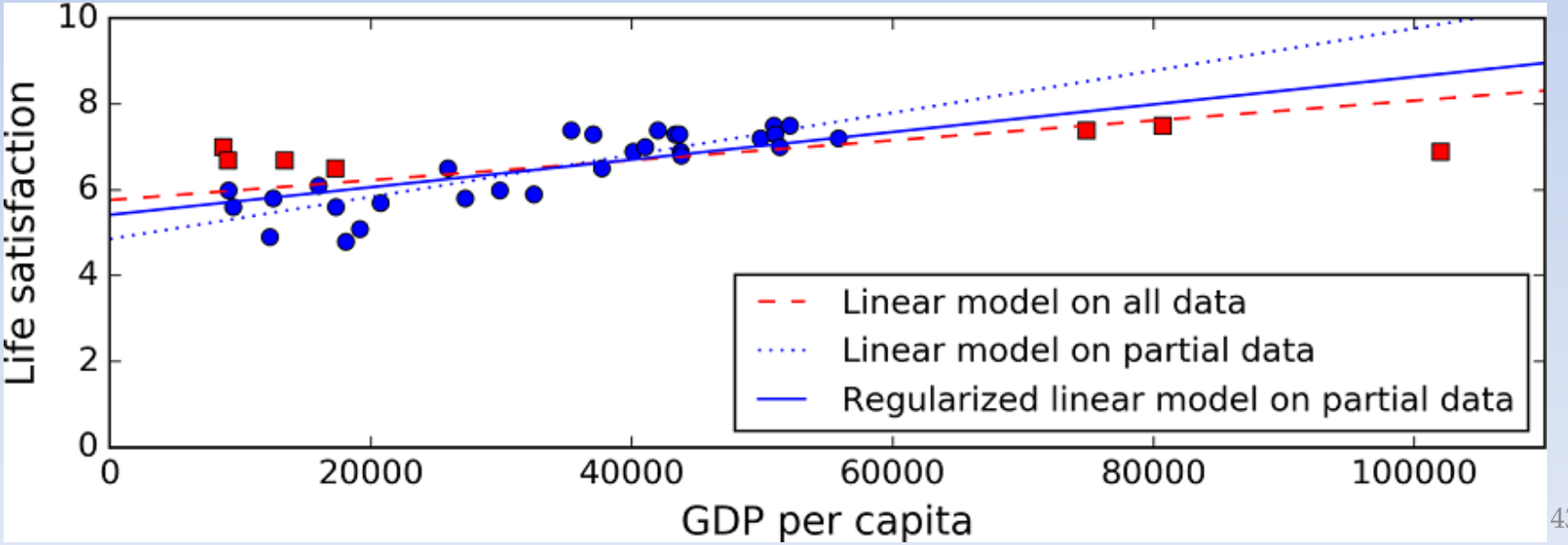
Challenges: Nonrepresentative Training Data

- In order to generalize well, it is crucial that your training data be representative of the new cases you want to generalize to.



Challenges: Overfitting/Underfitting





Challenges:

- Poor-Quality Data
- Irrelevant Features

Resources: Datasets

- Google Dataset Search:
  - https://datasetsearch.research.google.com/
- UCI Repository: http://www.ics.uci.edu/~mlearn/MLRepository.html
- Statlib: http://lib.stat.cmu.edu/
- Kaggle: http ://kaggle.com

Summary

- Machine Learning is about making machines get better at some task by learning from data, instead of having to explicitly code rules.
- There are many different types of ML systems
- In a ML project you gather data in a training set, and you feed the training set to a learning algorithm.
  - If the algorithm is model-based it tunes some parameters to fit the model to the training set, and then hopefully it will be able to make good predictions on new cases as well.
  - If the algorithm is instance-based, it just learns the examples by heart and uses a similarity measure to generalize to new instances.
- The system will not perform well if your training set is :
  - too small,
  - not representative, noisy, or polluted with irrelevant features (garbage in, garbage out).
- Lastly, your model needs to be neither too simple (in which case it will underfit) nor too complex (in which case it will overfit).