EMBEDDED VISION DESIGN 3

MACHINE LEARNING & DEEP LEARNING

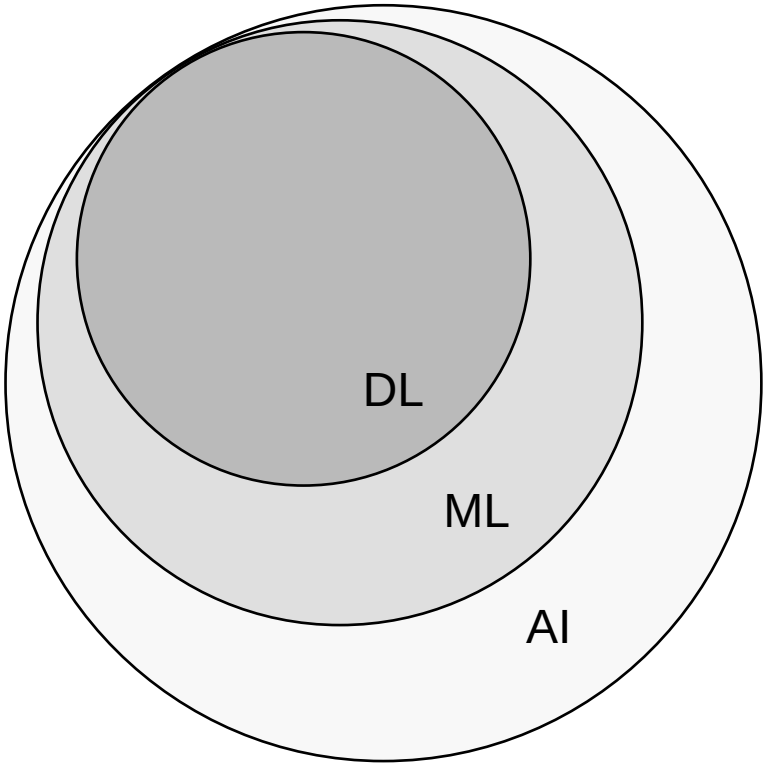
JEROEN VEEN

HAN_UNIVERSITY
OF APPLIED SCIENCES

CONTENTS

- Introduction
- Organization
- Why machine learning?
- Machine learning approaches
- Learning pipeline

DEFINING AI, DL & ML

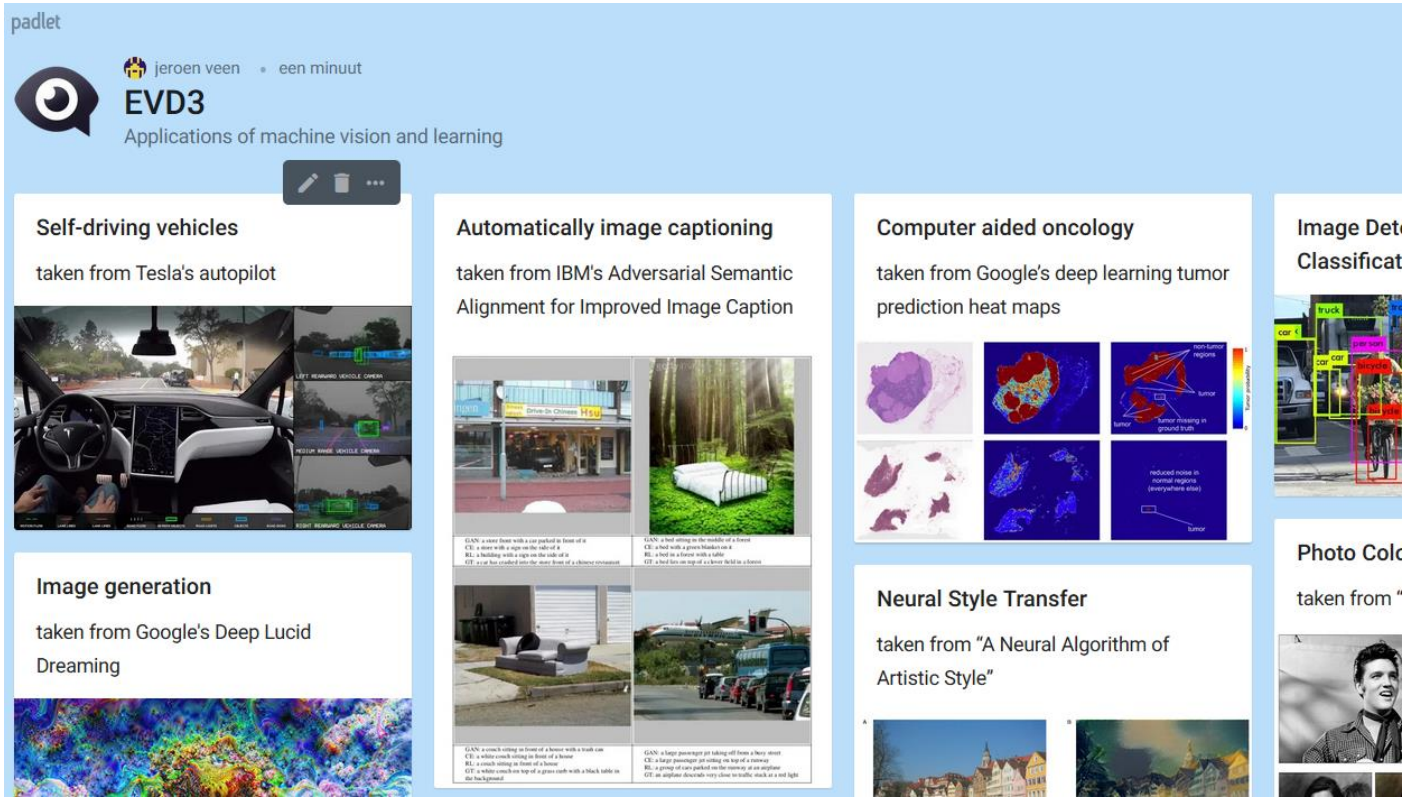


- Strong AI vs Applied AI
- Cognitive replication
- Rational process

Machine learning

- Performs predictive analysis
- Just fancy math & pattern matching

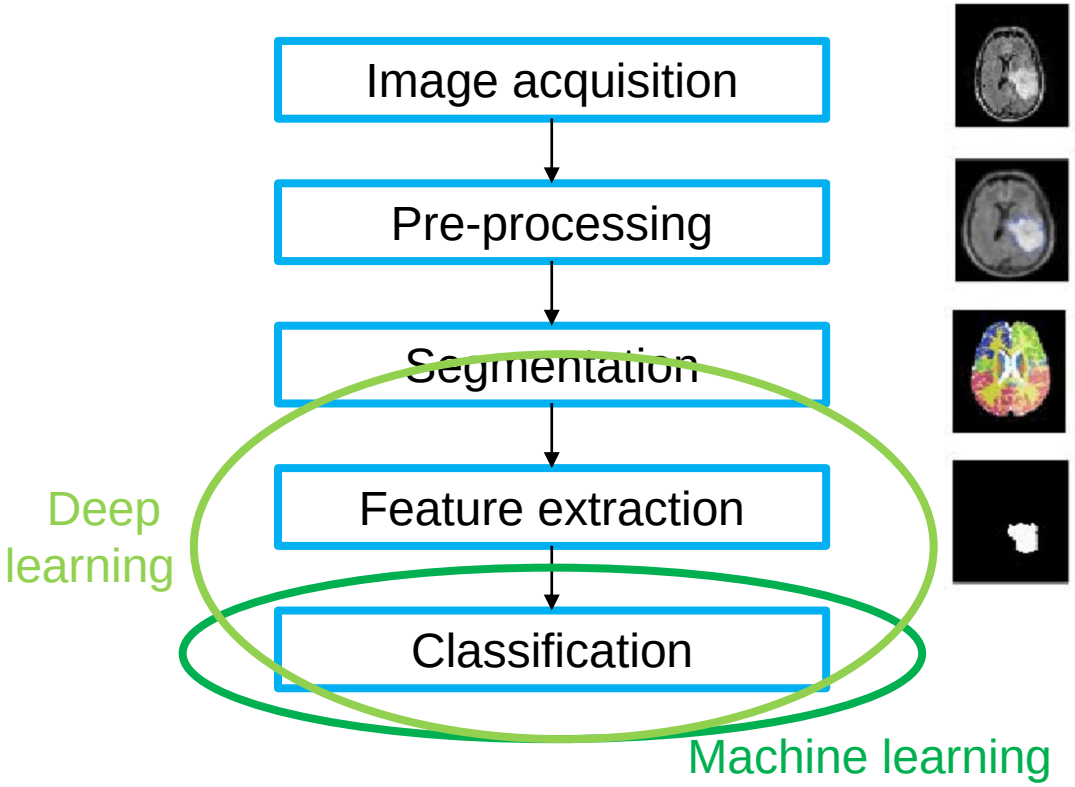
APPLICATIONS



https://padlet.com/jeroen_veen/zul8z8tbvhqpvb8t

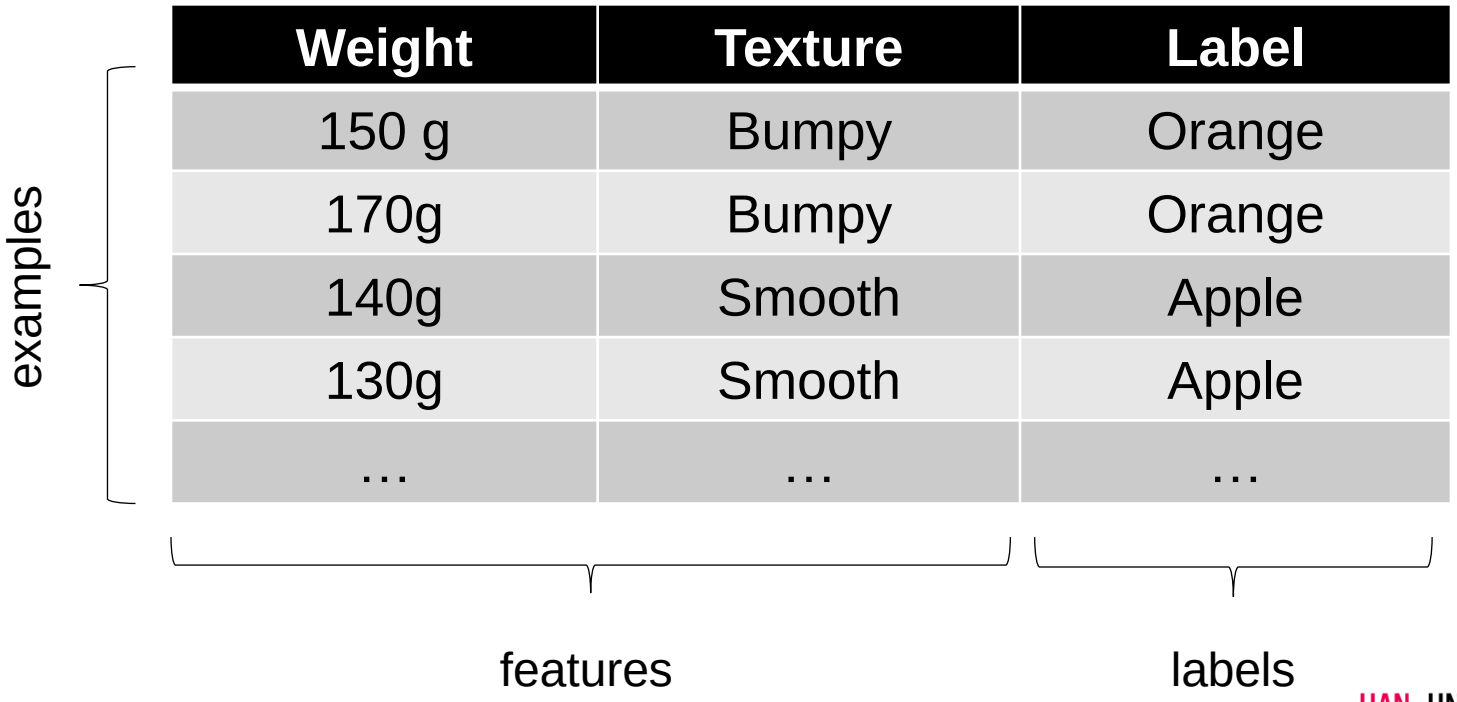
MACHINE LEARNING APPLIED TO VISION

- Classical image processing



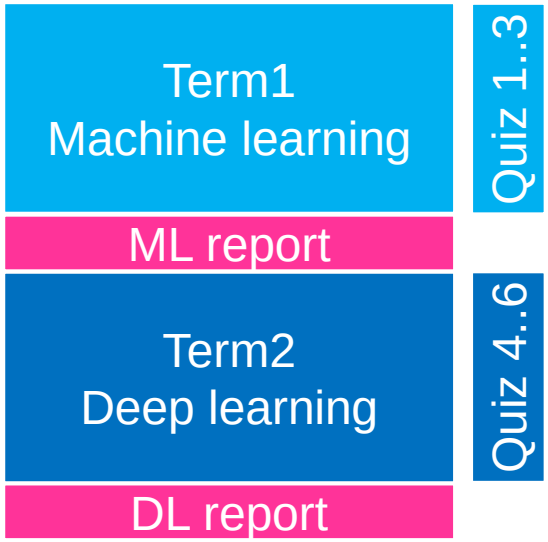
APPLES AND ORANGES

- https://www.youtube.com/watch?v=cKxRvEZd3Mw&feature=youtu.be



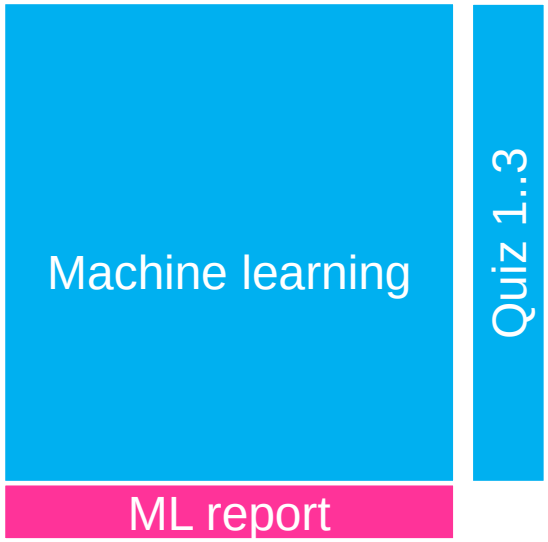
ORGANIZATION OF THE WORKSHOP (VT)

- Theory with integrated quizzes.
- Hands-on with 2 mini-projects
- Final mark:
  80% ML + DL report,
  20% quiz results
- Schedule on Gitlab



ORGANIZATION OF THE WORKSHOP (DT)

- Theory with integrated quizzes.
- Hands-on with mini-project
- Final mark:
  80% ML report,
  20% quiz results
- Live demo or short clip,
  showing your deployed model
- Schedule on Gitlab



QUIZ

- Individual, multiple choice questions
- Online: http://www.socrative.com room **1PTGB6PY**
- Open book quiz, so books and slides can be consulted

- **HAN student number**, so NOT your name, nickname or anything else.
- Quiz starts exactly at class hour and takes 10 minutes.
- Be on time and have your equipment prepared.
- During the quiz: no entering or leaving the classroom, and silence

QUIZ EXAMPLE

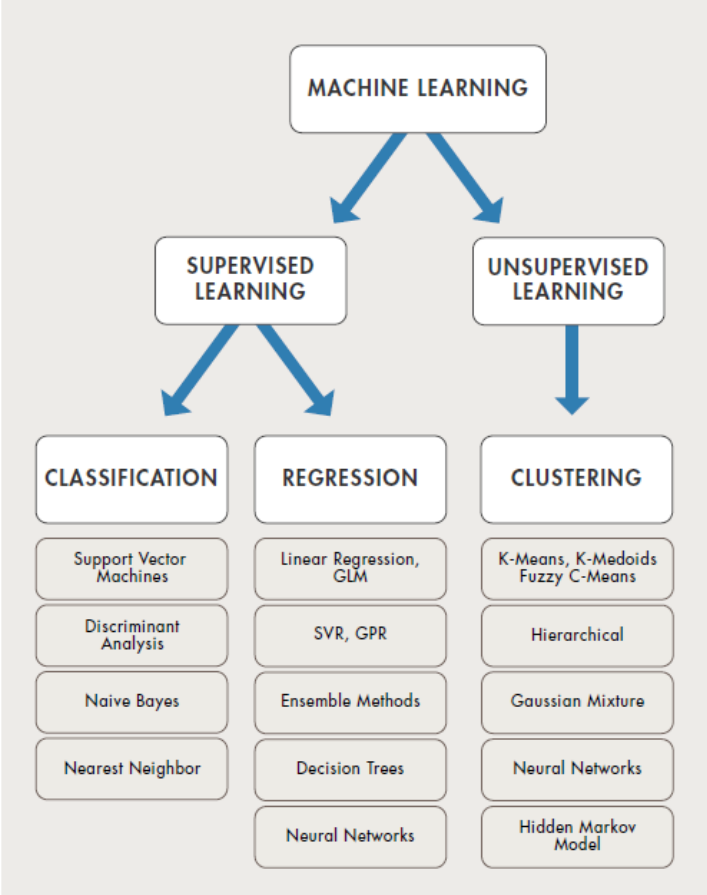
https://b.socrative.com/

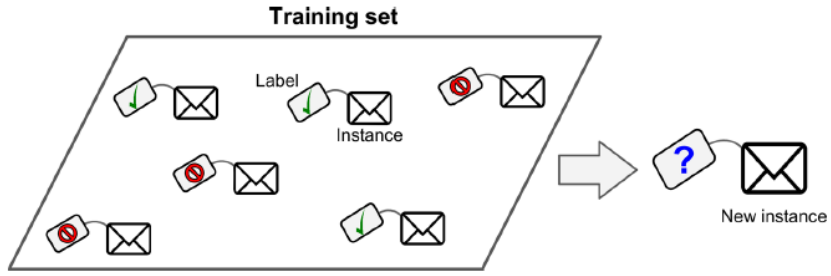
Room code: **1PTGB6PY**

EVD3 ASSIGNMENTS

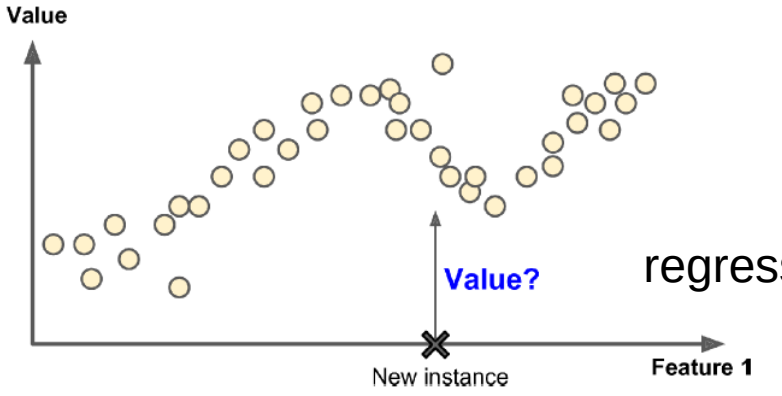
- A project team will consist of 3 students.
- Portfolio building using template
- Deliver intermediate results via HandIn
- Templates and schedule on Gitlab

MACHINE LEARNING APPROACHES





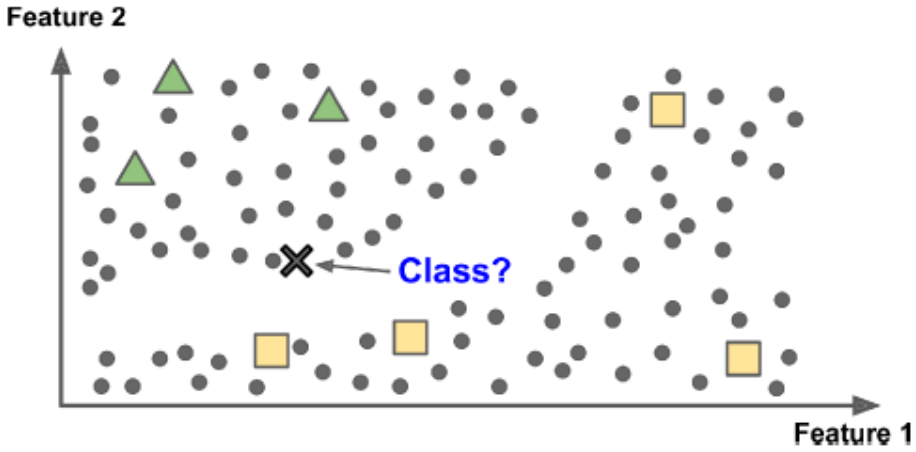
classification



regression

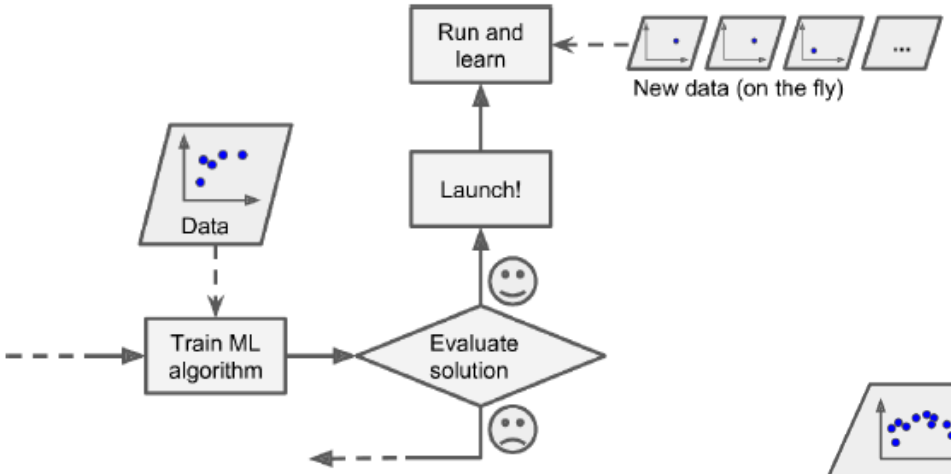
Source: Géron, ISBN: 9781492032632

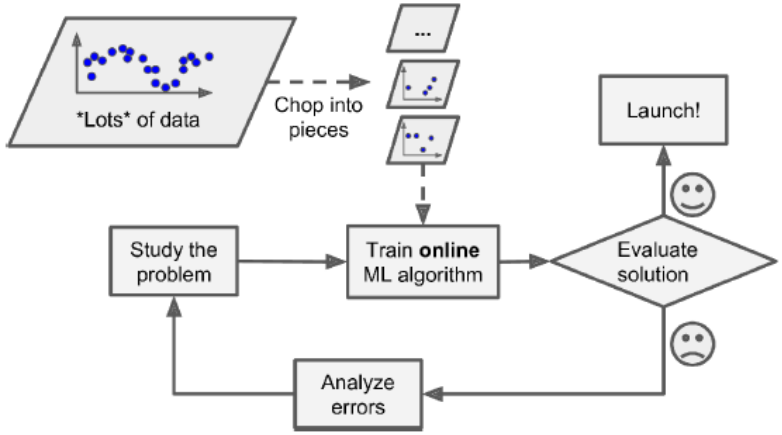
UNSUPERVISED AND SEMI-SUPERVISED LEARNING



Source: Géron, ISBN: 9781492032632

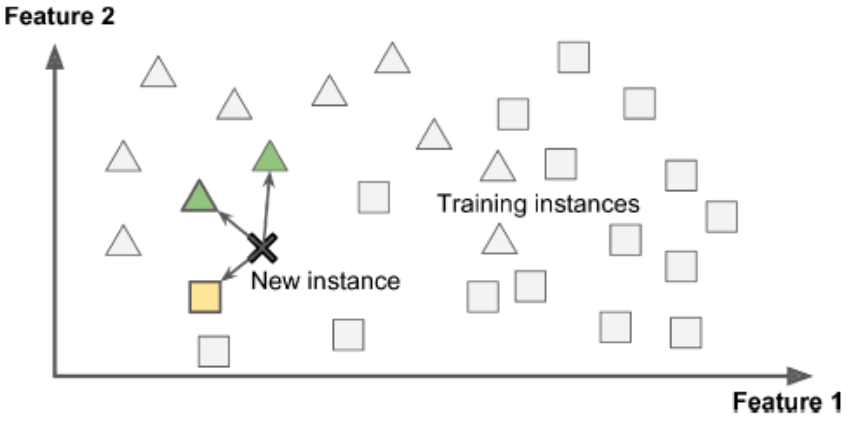
BATCH VS ONLINE LEARNING

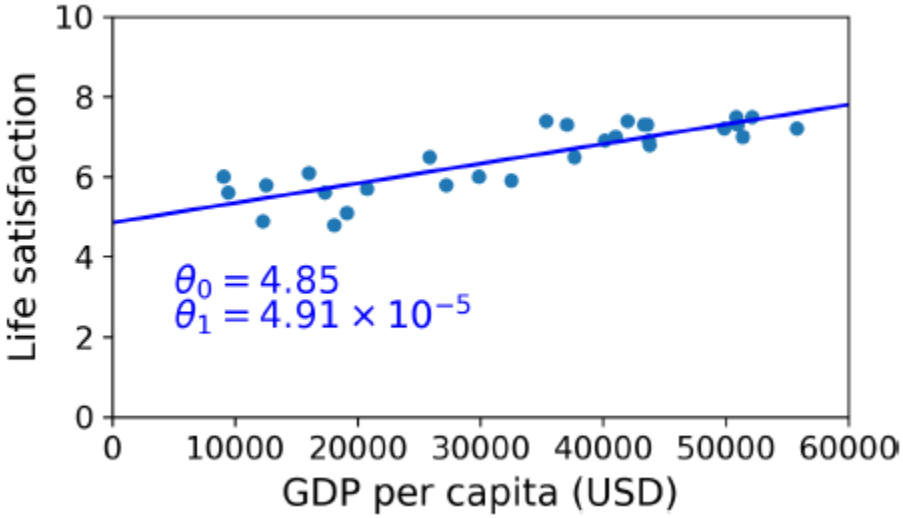




Source: Géron, ISBN: 9781492032632

INSTANCE-BASED VERSUS MODEL-BASED LEARNING



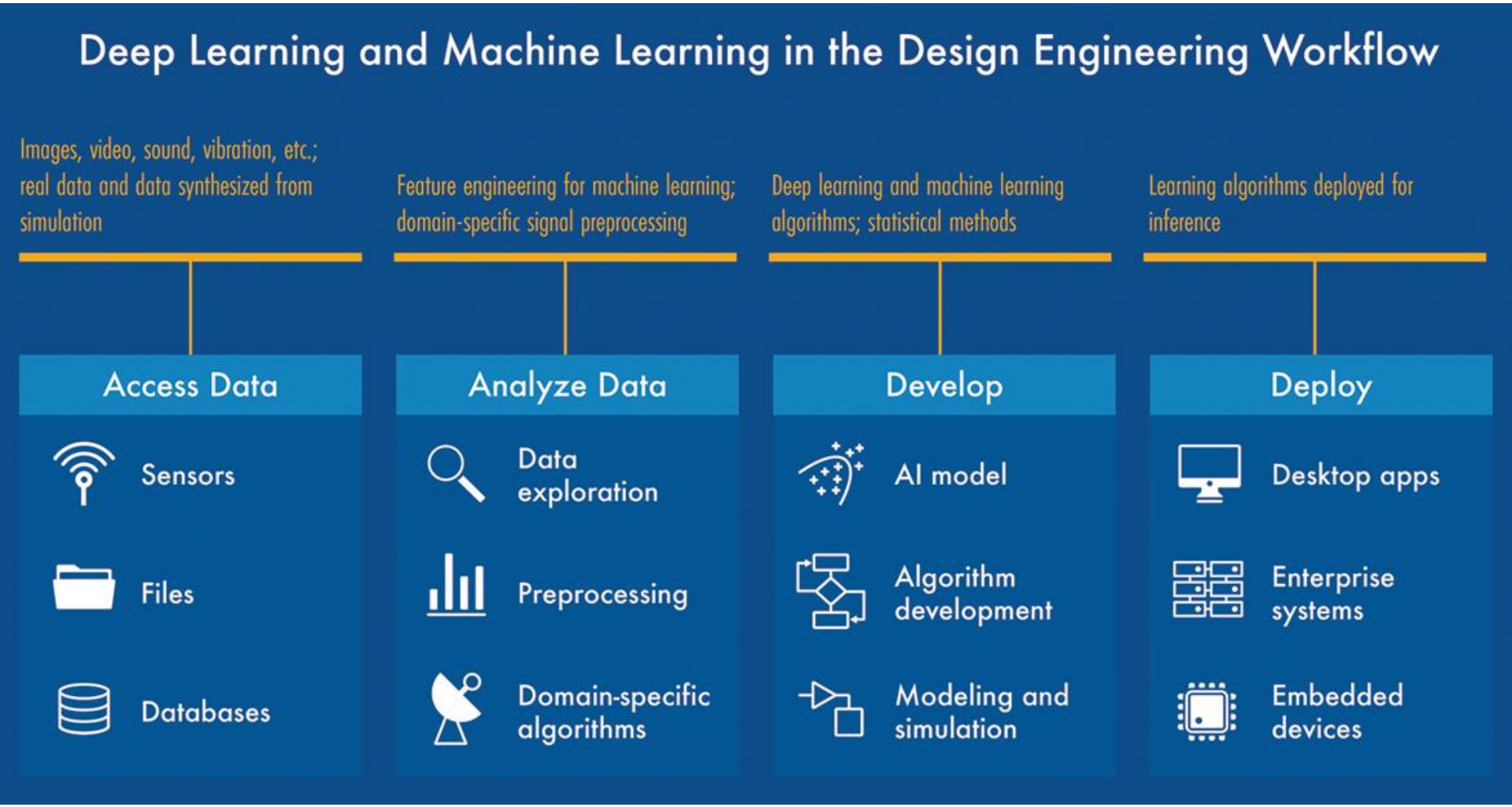


Source: Géron, ISBN: 9781492032632

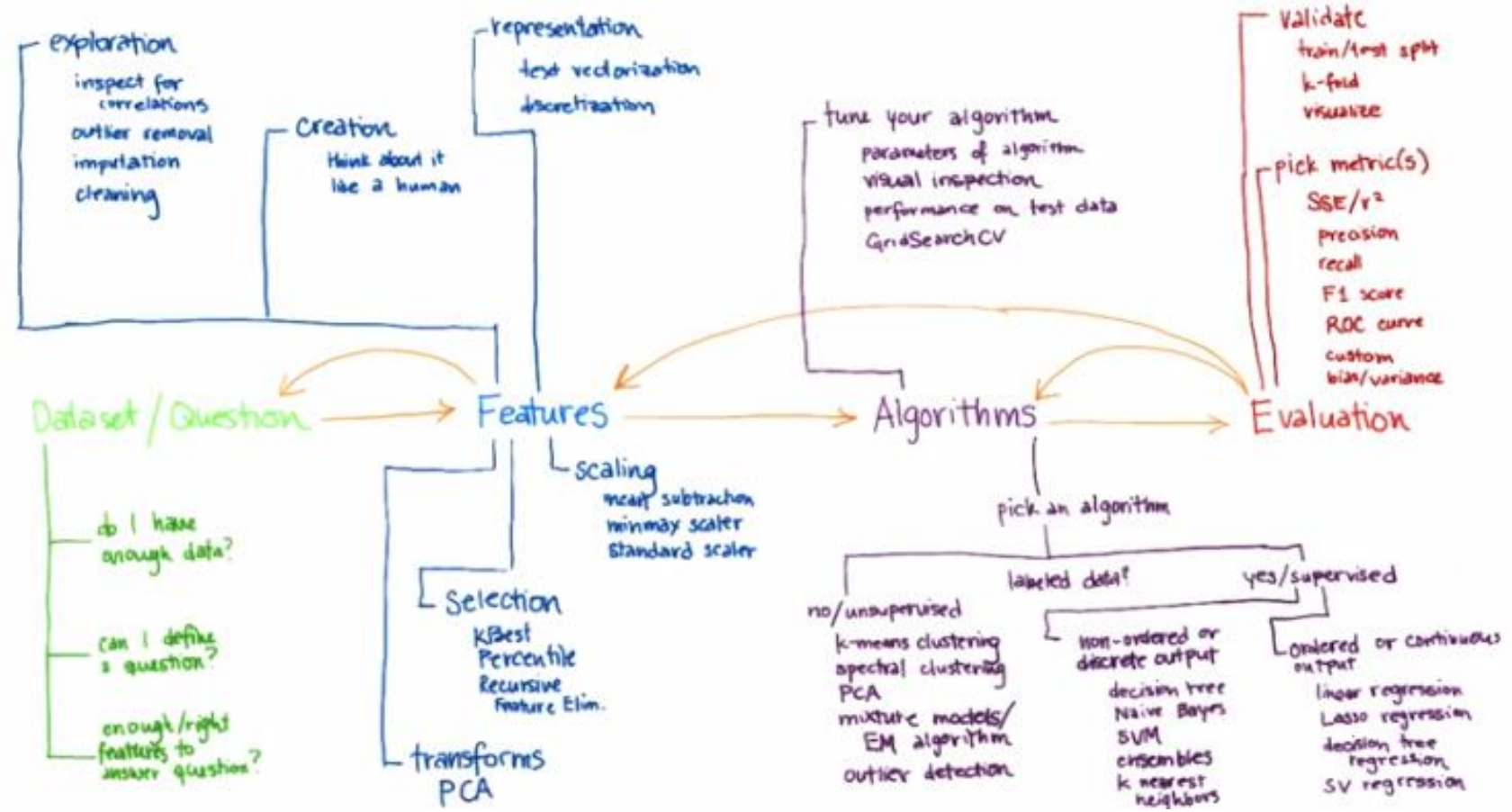
ML PITFALLS

- Massive amounts of training data is needed
- Labelling is tedious and error prone
- No relationship exists between input and output
- Solution is not transparent
- Solution fails to generalize
- Bias

WORKFLOW

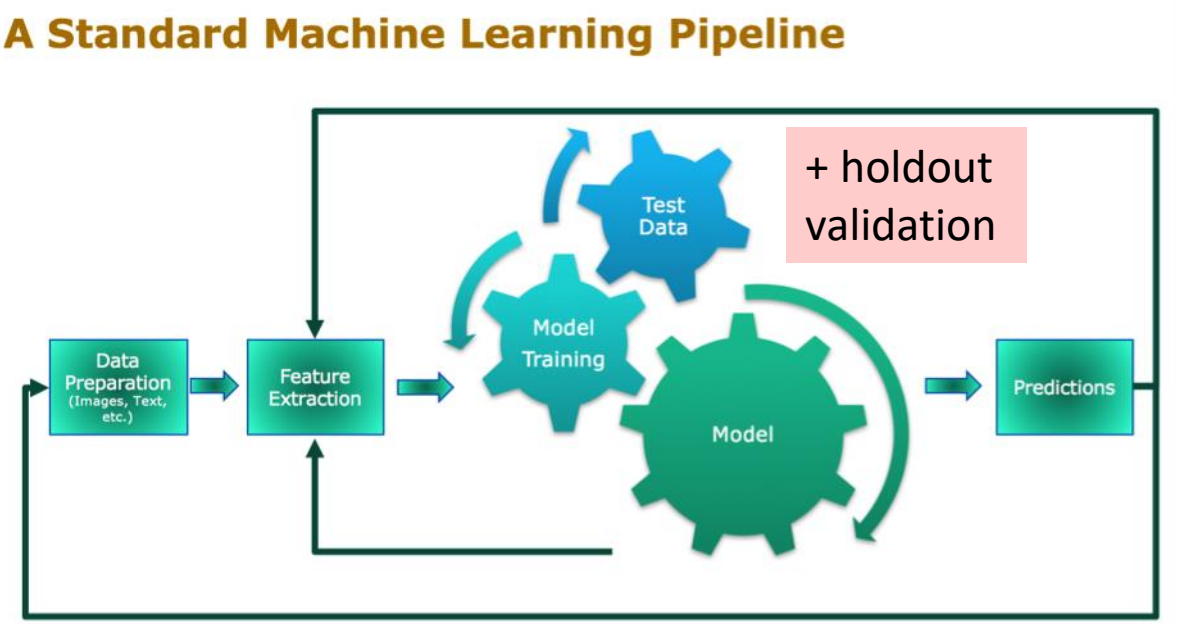

Deep Learning and Machine Learning in the Design Engineering Workflow
Images, video, sound, vibration, etc.; real data and data synthesized from simulation
Feature engineering for machine learning; domain-specific signal preprocessing
Deep learning and machine learning algorithms; statistical methods
Learning algorithms deployed for inference
Access Data
Sensors
Files
Databases
Analyze Data
Data exploration
Preprocessing
Domain-specific algorithms
Develop
AI model
Algorithm development
Modeling and simulation
Deploy
Desktop apps
Enterprise systems
Embedded devices

WORK FLOW



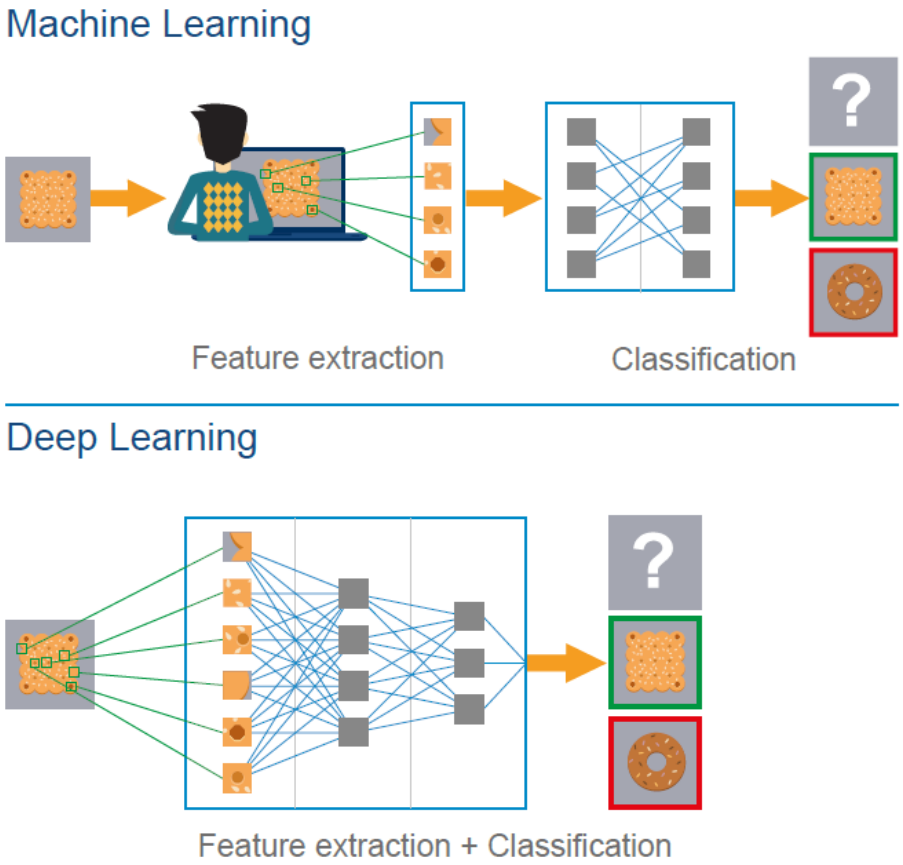
Source: Audacity

AUTOMATE MACHINE LEARNING WORKFLOW



Source: Western Digital

MACHINE LEARNING VS DEEP LEARNING



Source: Basler, Artificial Intelligence in Image Processing

ETHICS

- Self-adjustment can go horribly wrong
- Think of ‘sampling bias’, ‘exclusion bias’ and ‘prejudice bias’
- Context matters
- Transparency is becoming important
  General Data Protection Regulation (GDPR)

- **It is vital that developers take responsibility!**

OVERCOMING THE HYPE

- DL is not mature tech, lots of issues, constantly changing
- Not the only way of analysis and not the best
- E.g. self-driving cars combine with expert system
- You need sufficient and reliable data

	TECHNIQUE	DESCRIPTION
Anomaly Detection	Dynamic z-scores	Standard distribution measures are calculated for a given data set and uses a dynamic z-score threshold to detect anomalies
Leading/Lagging Indicators	Cross correlation	Measures with shared time series dimensions are analyzed to identify the time shift with the greatest correlation
Trend Lines	Regression analysis	Best fit line for time series data is estimated, and picks out the positive and negative trends that stand out the most
Data Segmentation	K-means clustering	Data points are recursively separated into logical groupings based on a set of local means

COMING UP - DATA

Weight	Texture	Label
150 g	Bumpy	Orange
170g	Bumpy	Orange
140g	Smooth	Apple
130g	Smooth	Apple
…	…	…

examples

features

labels

Pls read theory before next class

- Sampling noise, do you have sufficient data?
- Sampling bias, is your data representative?
- Data mismatch, is your data reliable?