Machine Learning Introduction

Jose Martinez Heras

01/03/2018

Resources

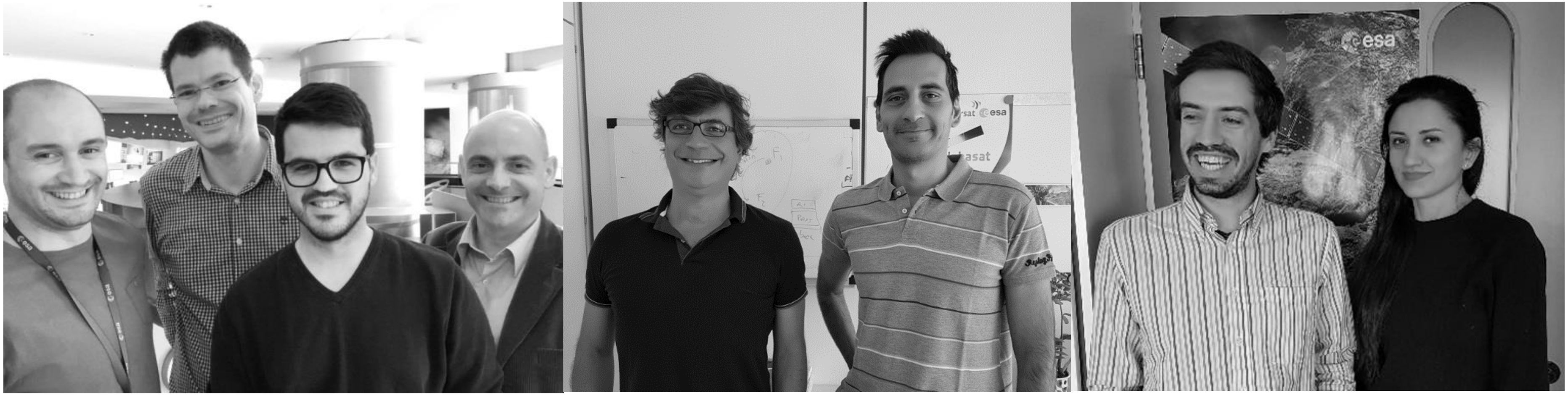
Watch the video of this lecture

https://dlmultimedia.esa.int/download/public/videos/2048/03/001/4803_001_AR_EN.mp4

Get presentation and additional resources on

https://github.com/jmartinezheras/2018-MachineLearning-Lectures-ESA

Advanced Mission Concepts & Technology



Redouane Boumghar, Jose Martinez-Heras, Jose Silva, Alessandro Donati, Simone Fratini, Nicola Policella, Rui Madeira, Lilli Bullinger

Guidelines for these lectures

Lectures

- Conceptual level
- Provide a intuition rather than diving into the maths
- Examples

Hands-on

- Work on examples to make the point of the lectures
- Will use python and numerical and machine learning libraries (free and open-source)

Lectures and Hands-on will be available on the Data Analytics ESA connect community

https://connect.esa.int/communities/community/data-analytics

Outline for Machine Learning Introduction

This session

- Introduction
- What is Machine Learning?
- Examples
- How Machine Learning works?
- Machine Learning concepts

Next Sessions

- Machine Learning techniques
- Hands-on

Artificial Intelligence

Reasoning

Machine Learning

Perception

Knowledge

Planning

Natural Language Processing

Move & manipulate objects

Artificial Intelligence

- Reasoning
- Knowledge
- Planning
- Machine Learning
- Natural Language Processing
- Perception
- Move & manipulate objects

Artificial Intelligence

Reasoning

Machine Learning

Perception

Knowledge

Planning

Natural Language Processing

Move & manipulate objects

What’s Machine Learning?

- **The science (and art) of programming computers so that they can learn from data [Aurélien Géron, 2017]**

- “The field of study that gives computers the ability to learn without being explicitly programmed” [Artur Samuel, 1959]

- “A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E” [Tom Mitchell, 1997]

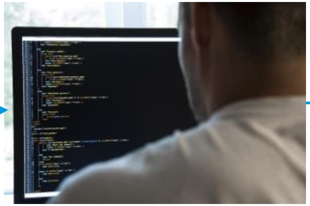
How is Machine Learning different from programming

Perform task

Data

Program

Results



Requirements

Program

Perform task

Data

Model

Results

Learning

Data

Results

ML

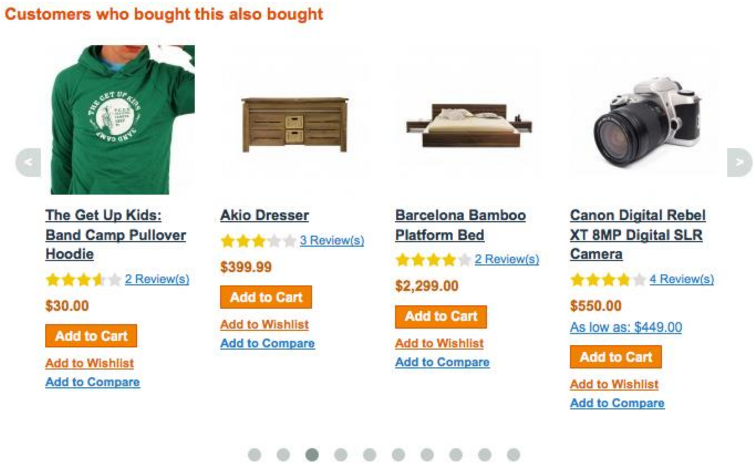
Model

Only for some tasks

Which tasks can Machine Learning perform?

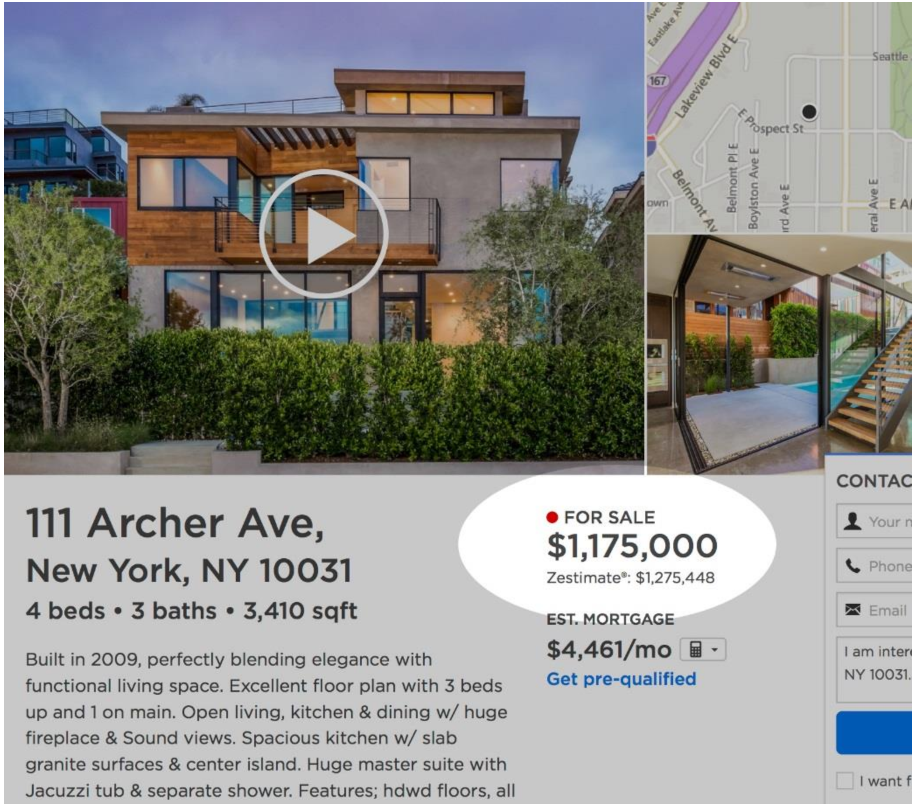
- Let’s see some examples

Spam Filter



Recommendation Engine

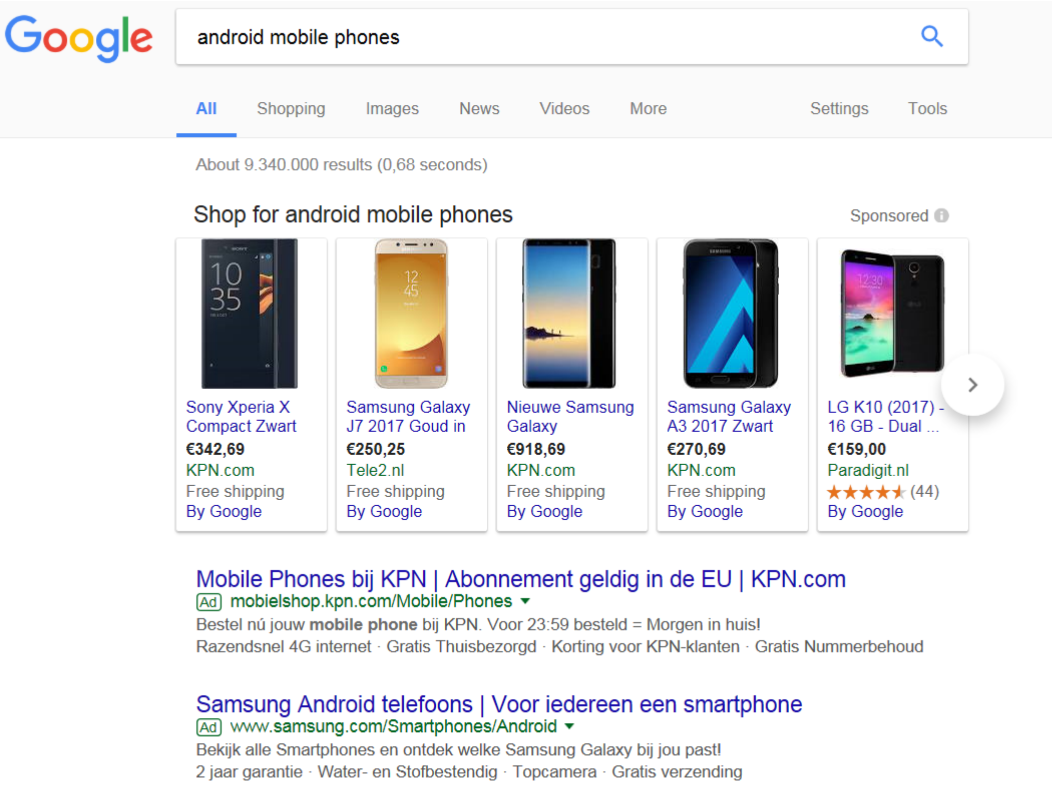
Which tasks can Machine Learning perform?



Real Estate

Predict at which price a property will be sold

Which tasks can Machine Learning perform?



Advertising

Predict which ads you are more likely to click on

Which tasks can Machine Learning perform?

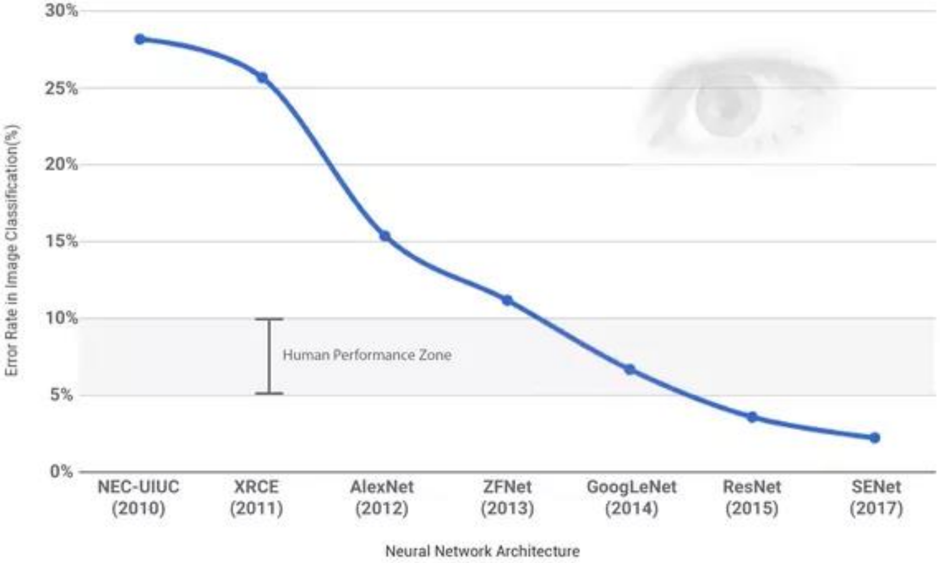
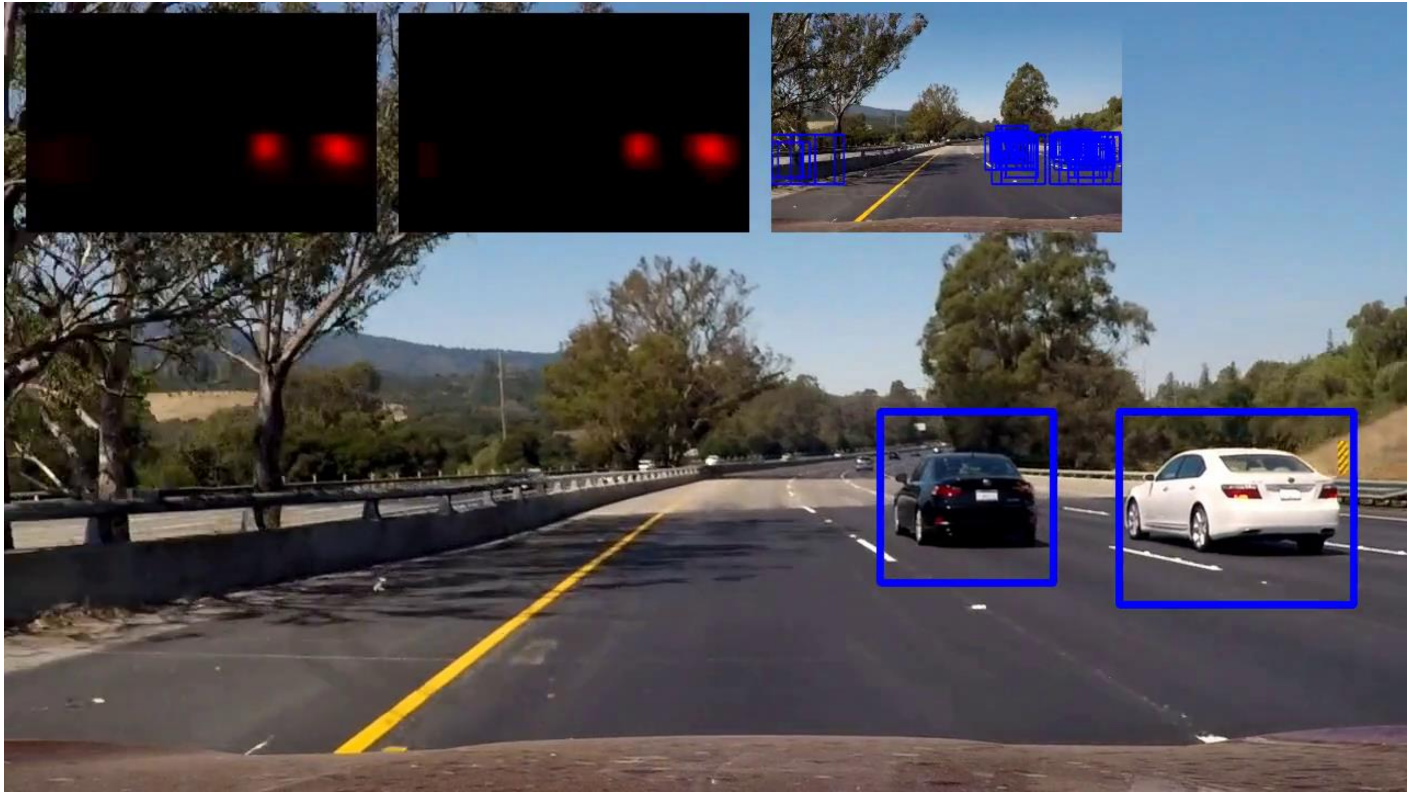


Image Classification

Which tasks can Machine Learning perform?



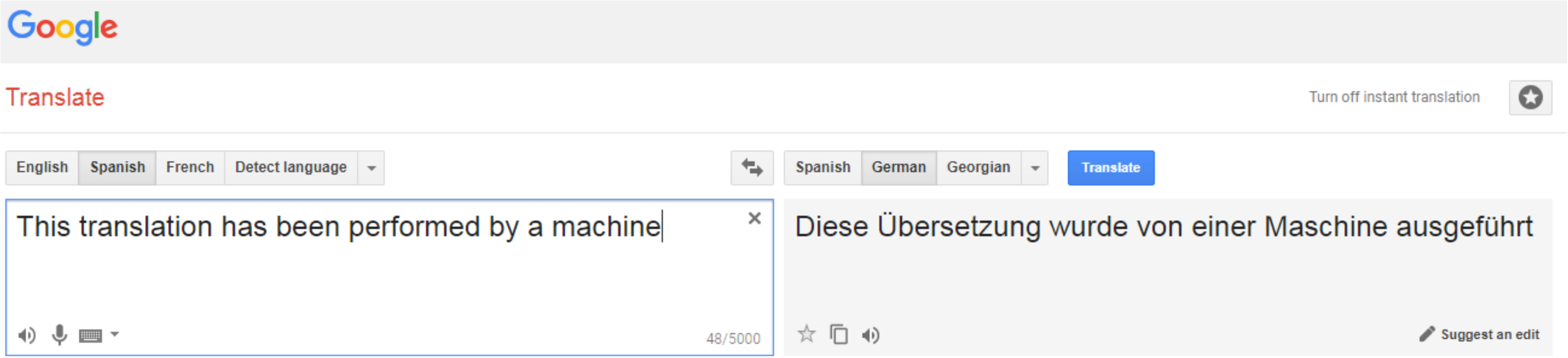
Self-driving cars

Which tasks can Machine Learning perform?

Speech Recognition
& Synthesis

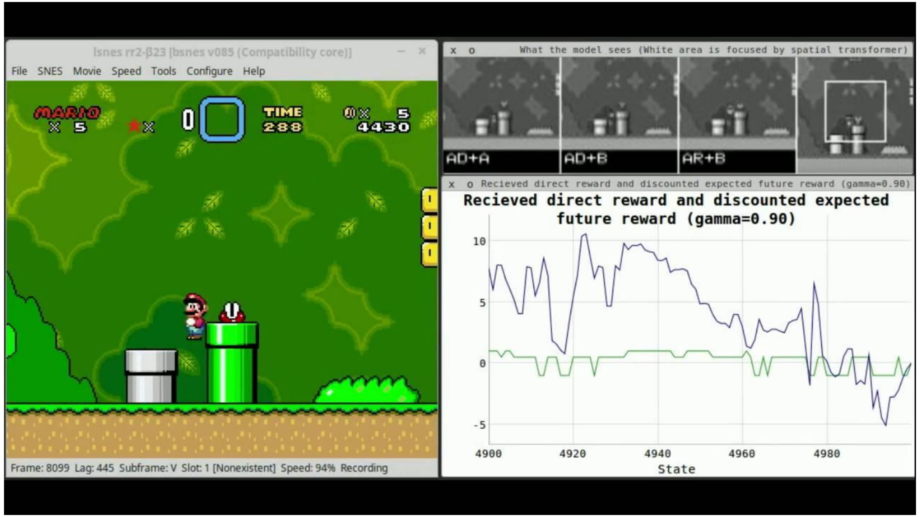
Which tasks can Machine Learning perform?

Language Translation



Which tasks can Machine Learning perform?

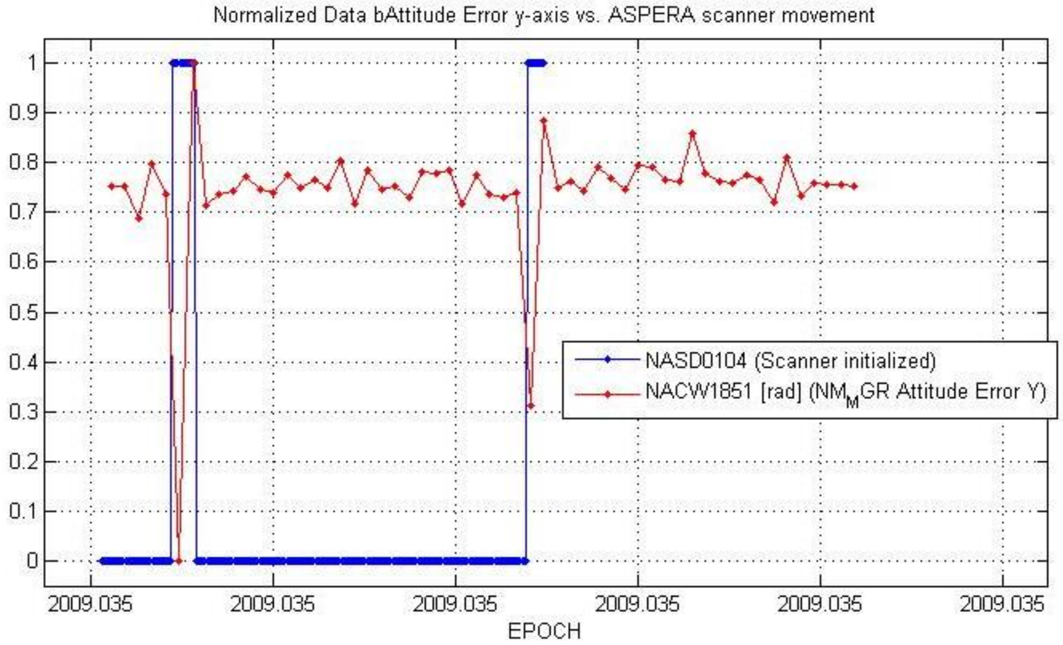
Playing Games



Some examples in Space Operations

Dr.MUST diagnostics: find the cause of an anomaly

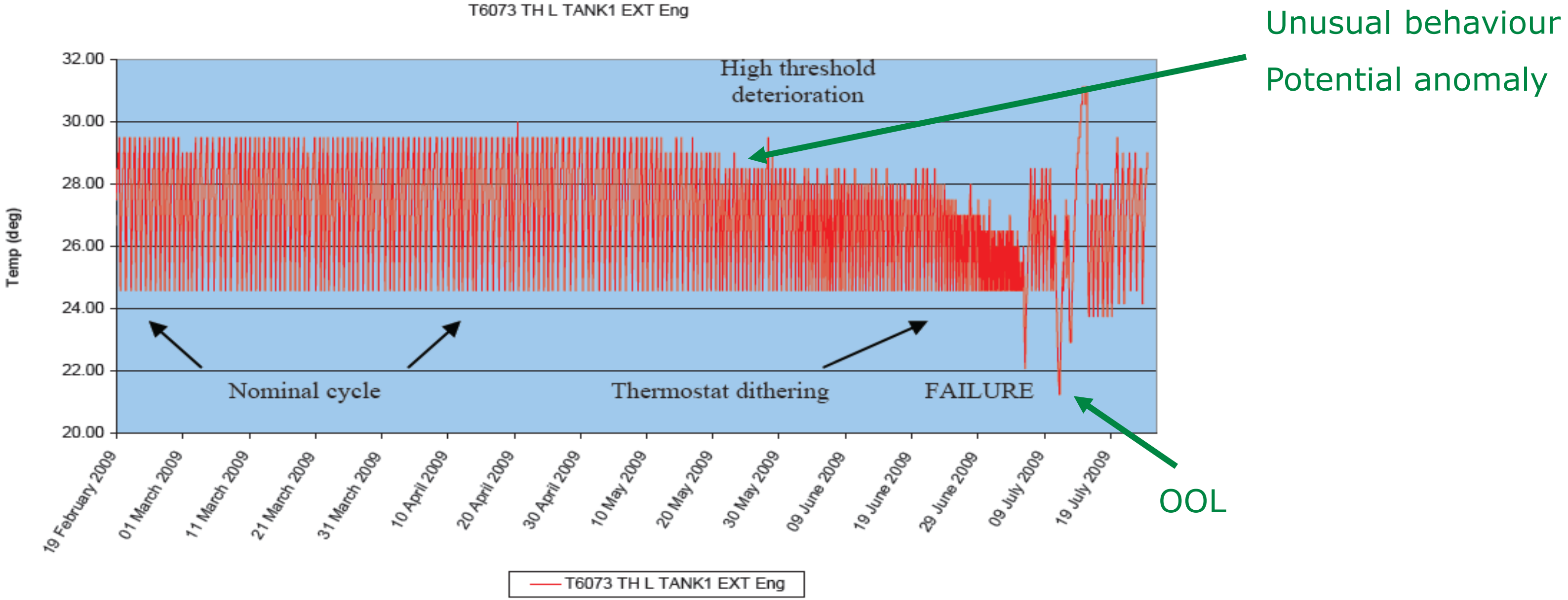
VEX attitude errors because of Aspera payload activation



Some examples in Space Operations

Novelty Detection: a novel behaviour is often the signature of an anomaly in the way to happen

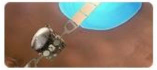
XMM Tank1 Ext temperature



Some examples in Space Operations

[MEX] Predict Thermal Power Consumption: Machine Learning competition

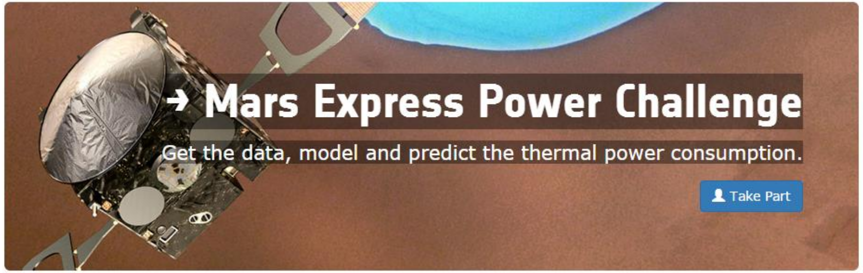
Kelvins | About | Competitions | Mars Express Power Challenge | jose



- Home
- Challenge
- Data
- Evaluation
- Rules
- Results
- Leaderboard
- Submission
- Discussion

March 31, 2016, 10 p.m. UTC — Timeline — July 31, 2016, 10 p.m. UTC

The competition is over.



It has now been more than 12 years that the Mars Express Orbiter (MEX) provides science data from Mars about its ionosphere and ground subsurface composition. The 3D imagery of Mars has provided the community with unprecedented information about the planet. Today, thanks to the work of careful and expert operators, Mars Express Orbiter still provides information that supports ground exploration missions on Mars (Curiosity, Opportunity, ...) and a lot of other research.

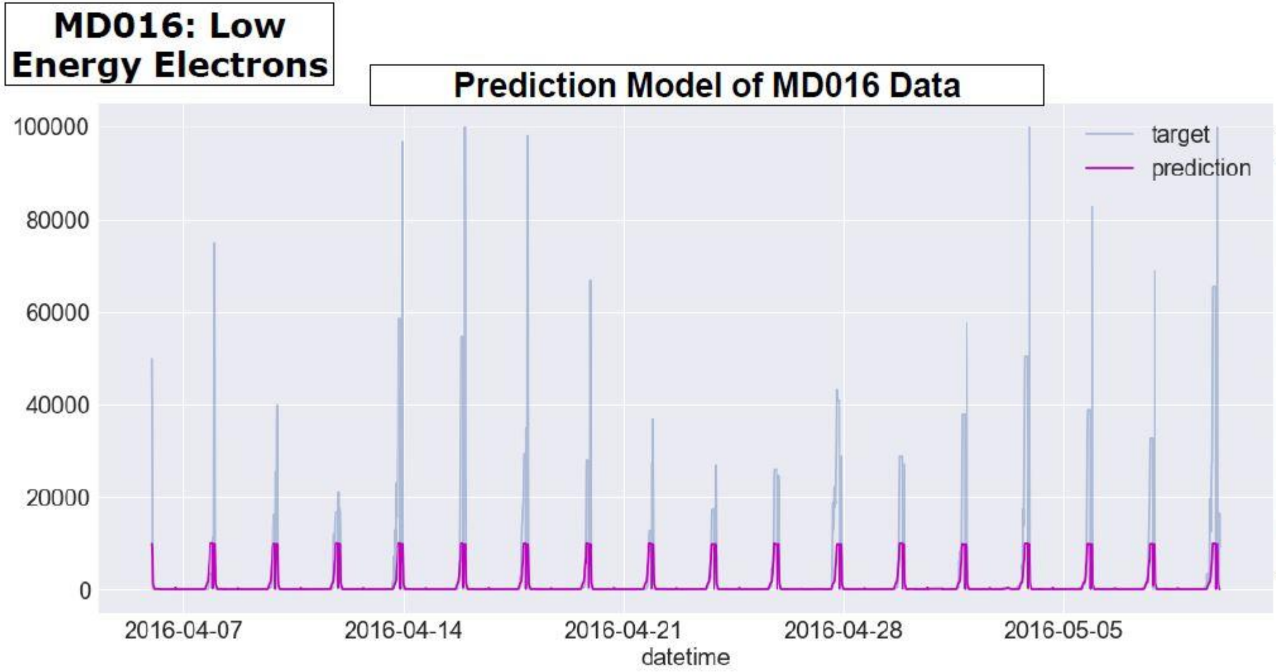
The Mars Express Orbiter is operated by the European Space Agency from its operations centre (Darmstadt, Germany) where all the telemetry is analysed. The health status of the spacecraft is carefully monitored to plan future science observations and to avoid power shortages.

Operators of Mars Express keep track of the thermal power consumption thanks to the telemetry data. The spacecraft uses electric power coming from the solar arrays (or batteries, during eclipses) not only to supply power to the platform units, but also to the thermal subsystem, which keeps the entire spacecraft within its operating temperature range. The remaining available power can be used by the payloads to do science operations:

Science Power = Produced Power - Platform Power - Thermal Power

Some examples in Space Operations

[Integral] Radiation Belts entry / exit times prediction to increase science return



Importance of features

0.820534935602, altitude
0.0507381593424, rx
0.0476677808188, vx
0.0302844234954, ry
0.0129121849191, vz
0.0106376974037, rz
0.0105089758999, vy
0.00720441182158, PA_rad
0.00315537128504, DEC_rad
0.00272178660137, RA_rad
0.00222400982, ROLL_ANGLE
0.00137769977878, umbra
3.41720505552e-05, penumbra

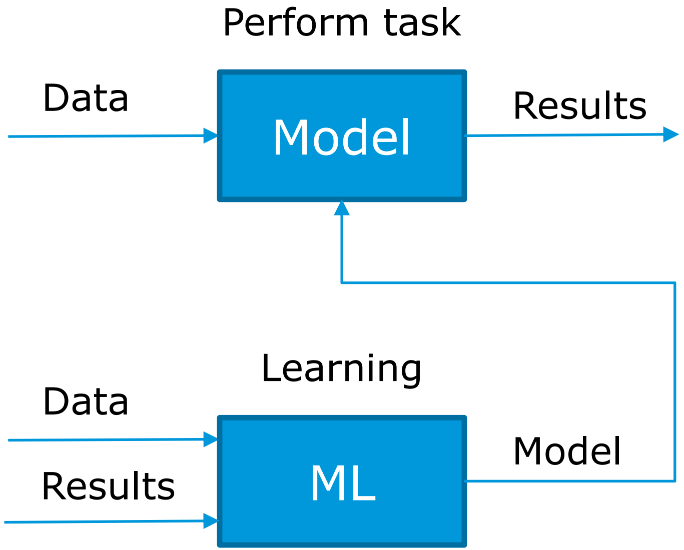
Some examples in Space Operations

[esa.int] Predict which articles will receive a high number of views

High: Rosetta, comet, surface, lander, image, crater, mars, stars, galaxy, black holes

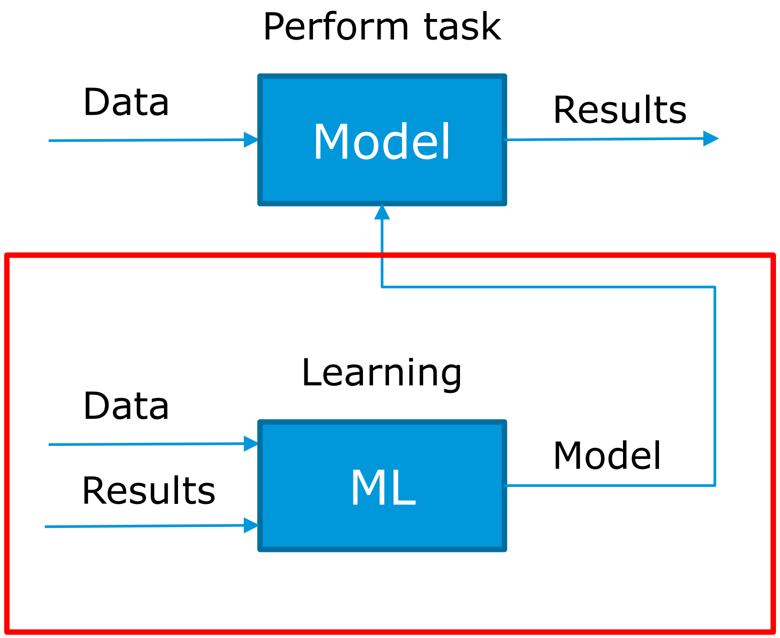
Low: ESA, company, technology

Considerations on using Machine Learning



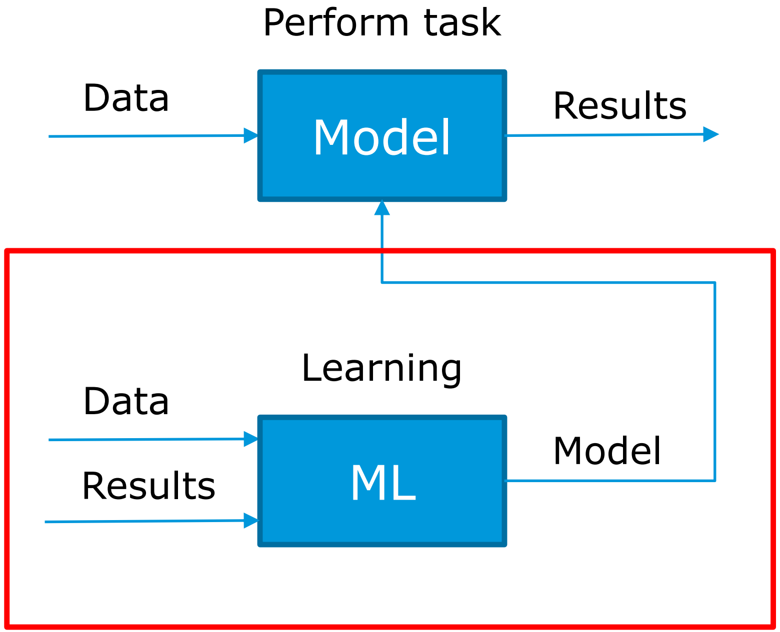
- There must be a pattern in the input – output relationship (lottery winning numbers cannot be predicted with ML)
- There must be enough data to discover this pattern
- It’s difficult to formulate a mathematical expression (otherwise we will just use this formula instead)

How Machine Learning Learns?



- Different Machine Learning techniques for different kinds of tasks
- Learning is finding which model's parameters represent best the input – output mapping

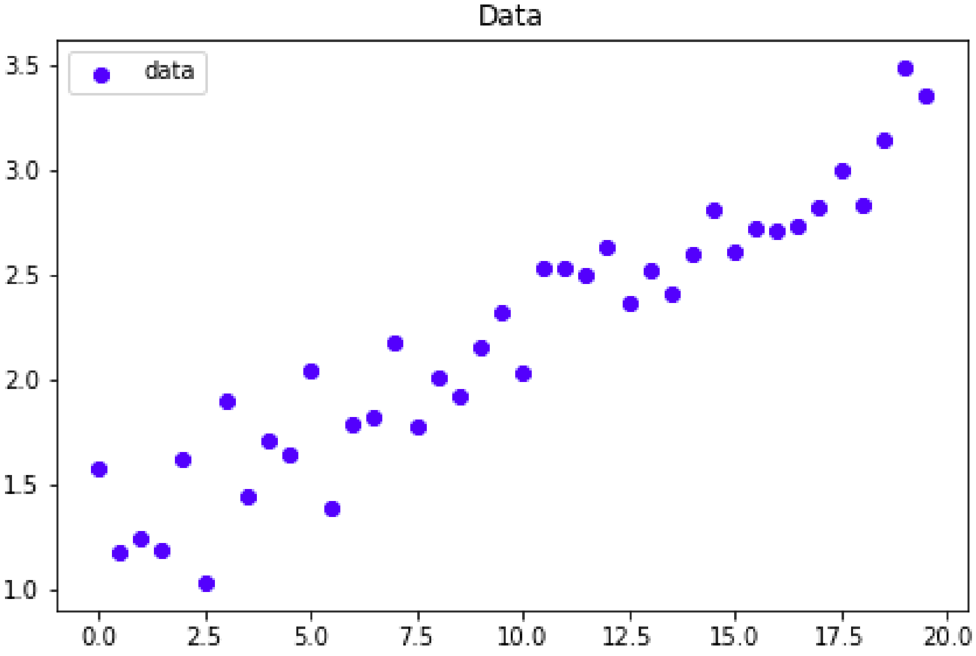
How Machine Learning Learns?



- Linear Regression example:
  - Model: f(x) = mx + b
  - Model's parameters: m, b
  - Parameter values: m=1, b=0

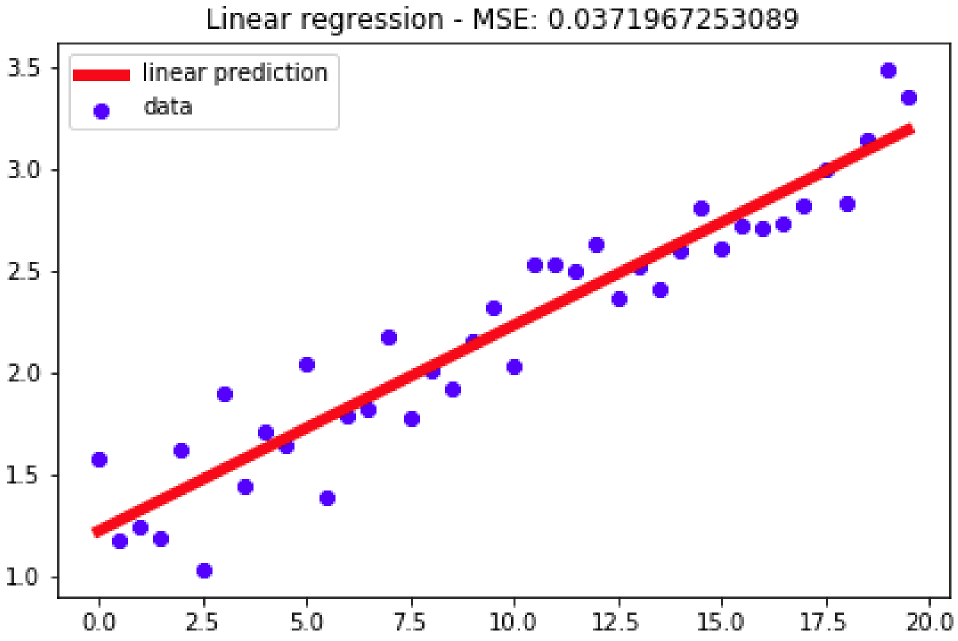
- Learning is finding which values of 'm' and 'b' fit the data best (e.g. minimizes the prediction error)

How Machine Learning Learns?



Find the model parameters values that minimize the error

How Machine Learning Learns?

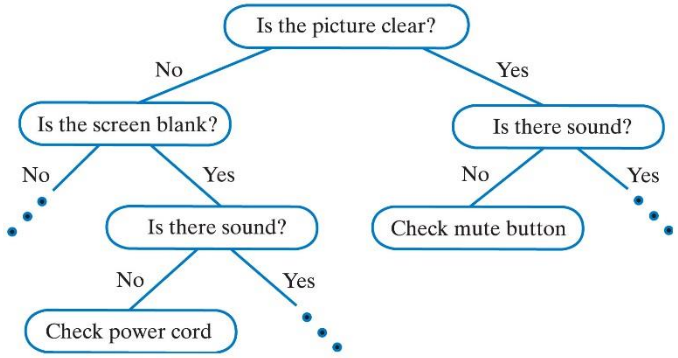


Find the model parameters values that minimize the error

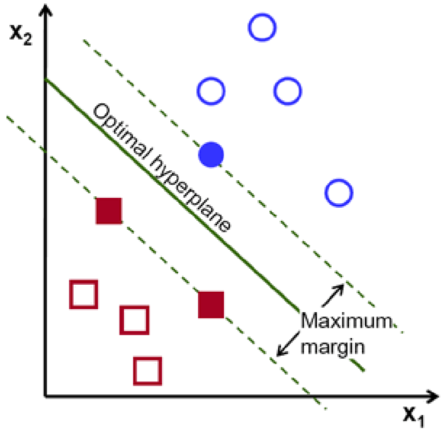
m = 0.1014
b = 1.2258

How Machine Learning Learns? – some models

Polynomials: $ax + bx^2 + \ldots + c$



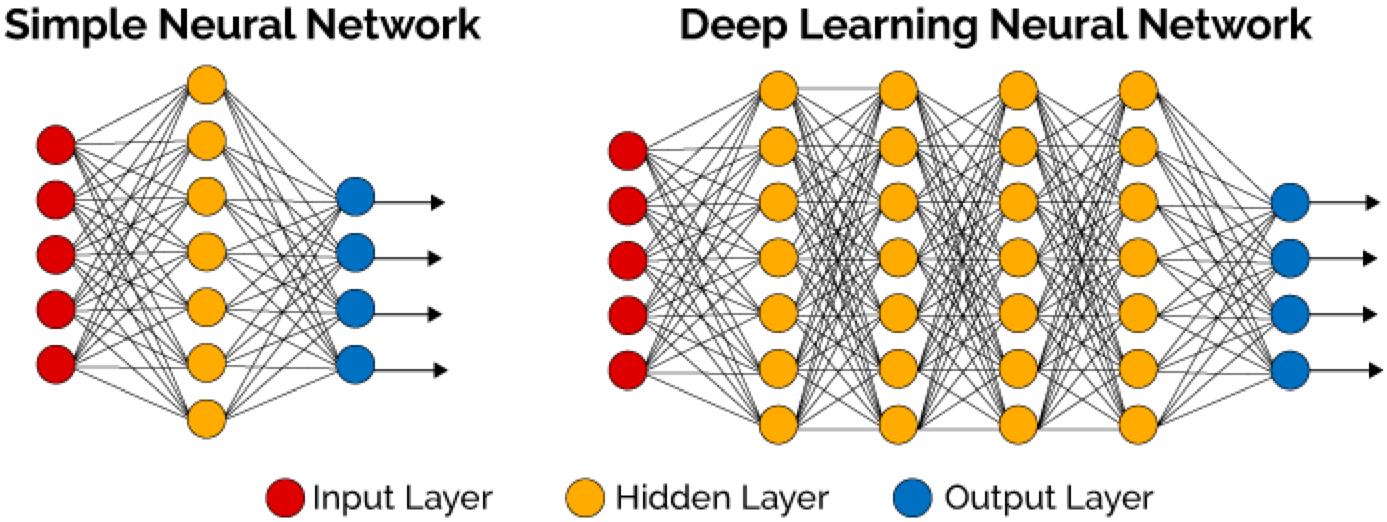
Decision trees: which nodes, which decisions



Support Vector Machines (SVM): Vectors

How Machine Learning Learns? – some models

Neural Networks / Deep Learning: weights



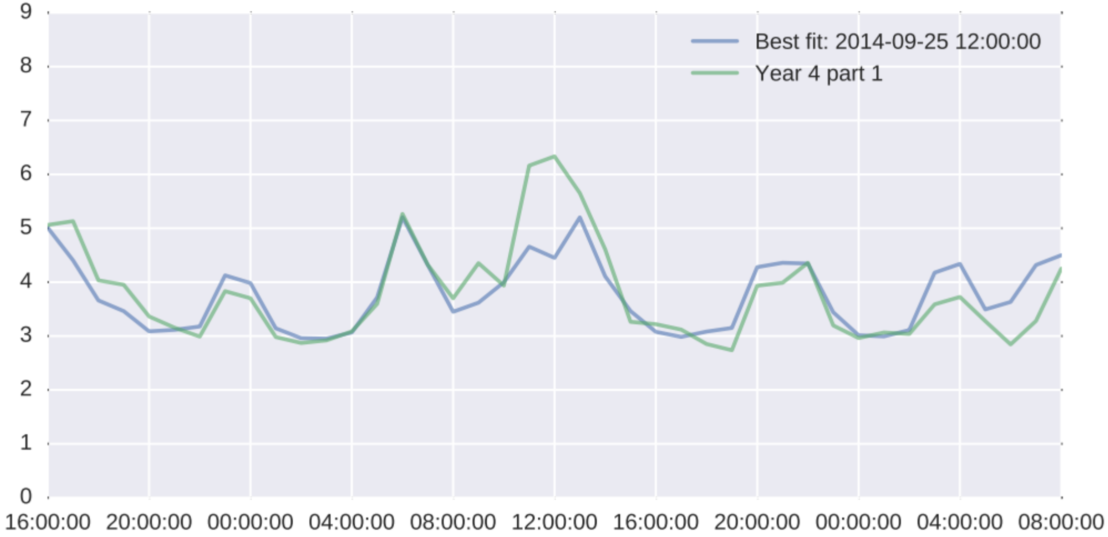
Types of Learning

Depending on the level of supervision …

- Supervised
- Unsupervised
- Semi-supervised
- Reinforcement Learning

Types of Learning

- Supervised
  - Supervision: we can tell for every case what the correct answer was
  - Example: predict the thermal power consumption



Focus
Predict the future

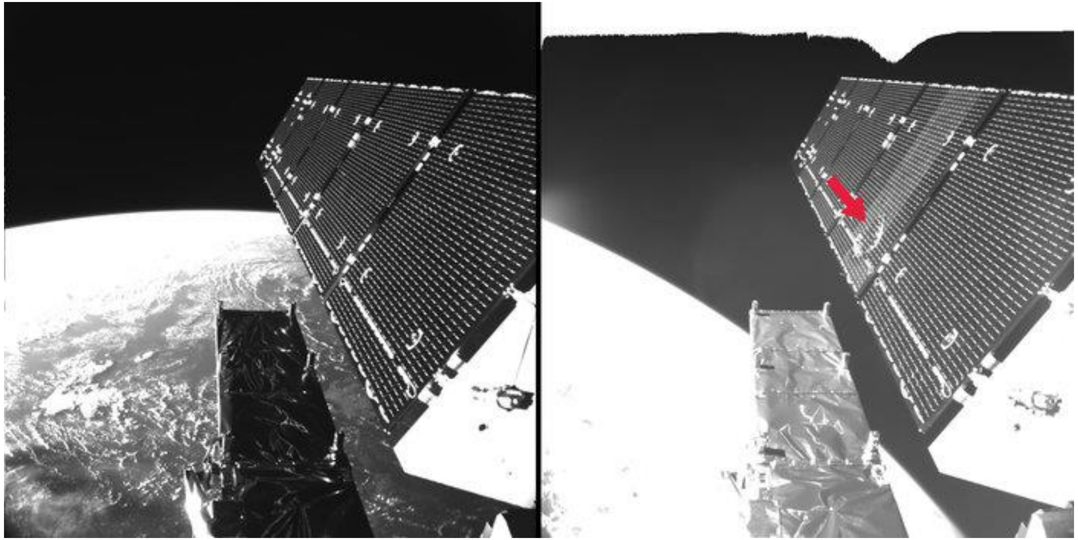
Types of Learning

- Unsupervised
  - Supervision: there is no right answer, we are looking for insights
  - Example: market basket analysis for supermarkets

Focus
Understand the past

Types of Learning

- Semi-supervised
  - Supervision: we can tell the correct output for a limited number of cases
  - Example: characterize what a particle impact looks like in TM



Focus
Understand the past

Sentinel-1A: particle impact on August 23th 2016

Types of Learning

- Reinforcement Learning
  - Supervision: we only know the final outcome, but not intermediate steps
  - Example: playing Go

Focus
Find which is the next action most likely to lead to the desired outcome

Types of Learning

- The type of learning with most industrial applications is

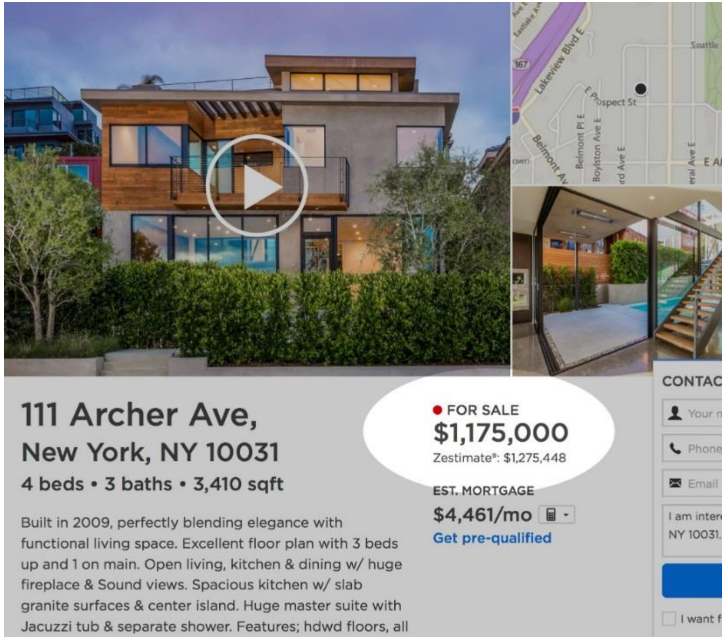
Supervised Learning

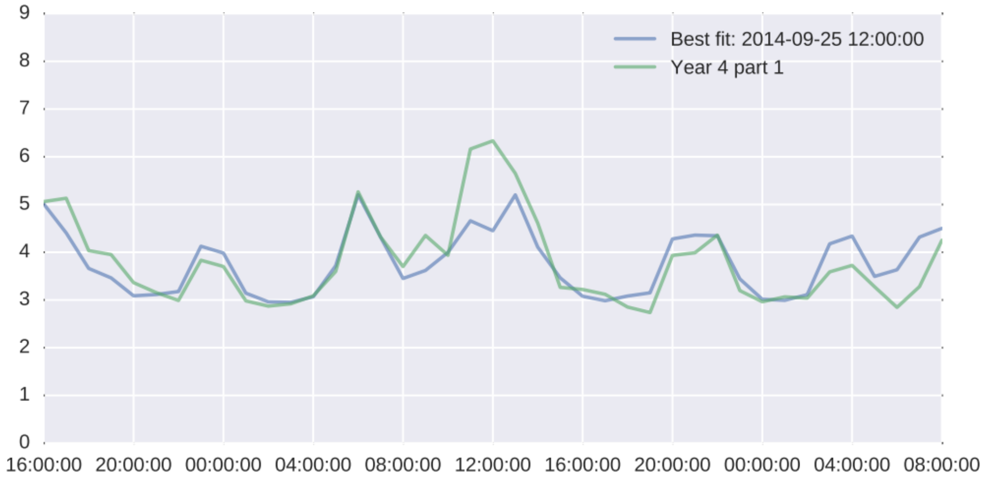
(Predictive Analytics)

- Depending of what kind of data is predicted we can talk about:
  - Regression
  - Classification

Regression

- Predict real numbers





Classification

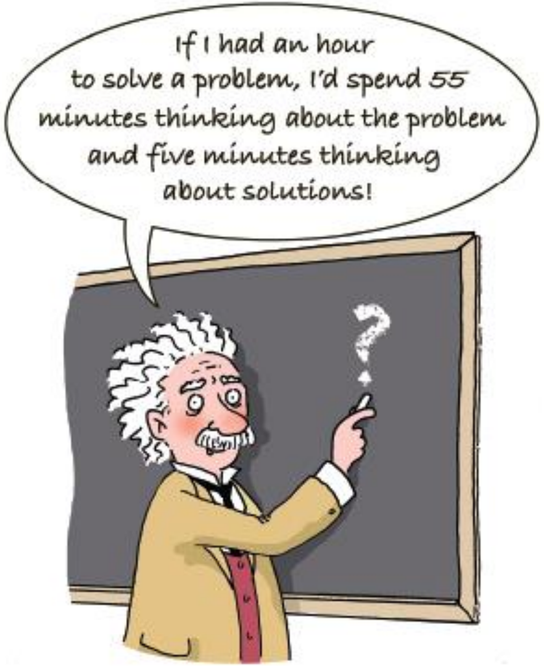
- Predict which option out of a limited set of possibilities

Spam Filter

What’s in the picture

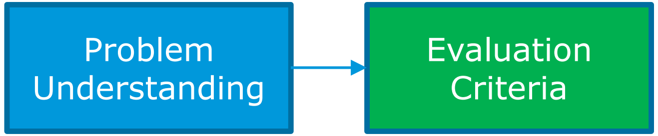
Machine Learning Workflow

Problem Understanding



Machine Learning Workflow

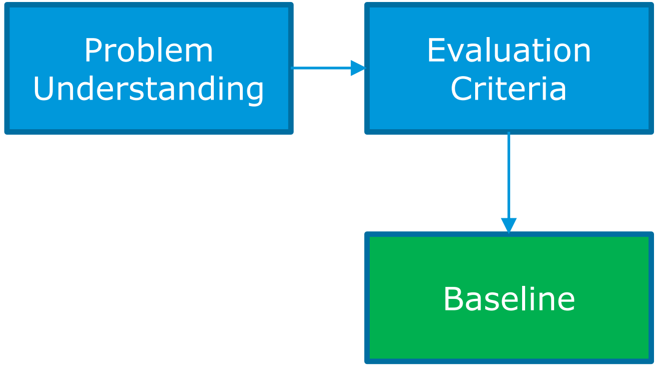
- How will we measure how good the model is performing?
- Do we know already at what point it would be enough?



Machine Learning Workflow

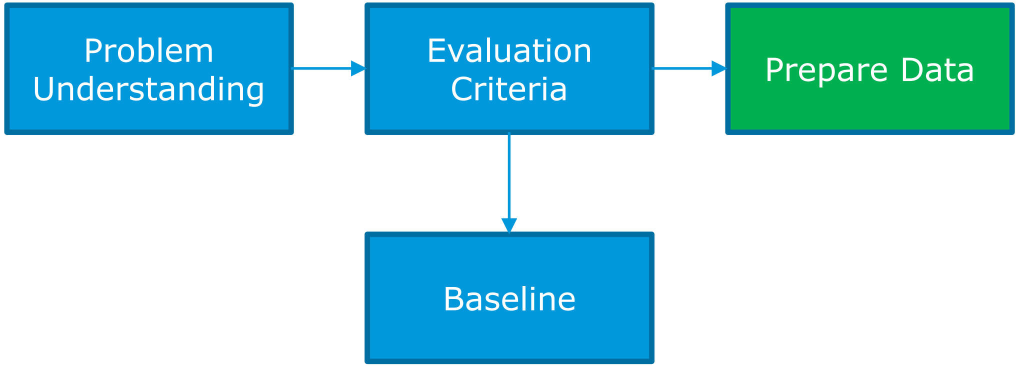
- How the current approach is performing against the evaluation criteria? Define a simple baseline if there is none (e.g. mean value)

- This will allow us to quantify how much Machine Learning helps and if it is worthwhile compared to simpler solutions



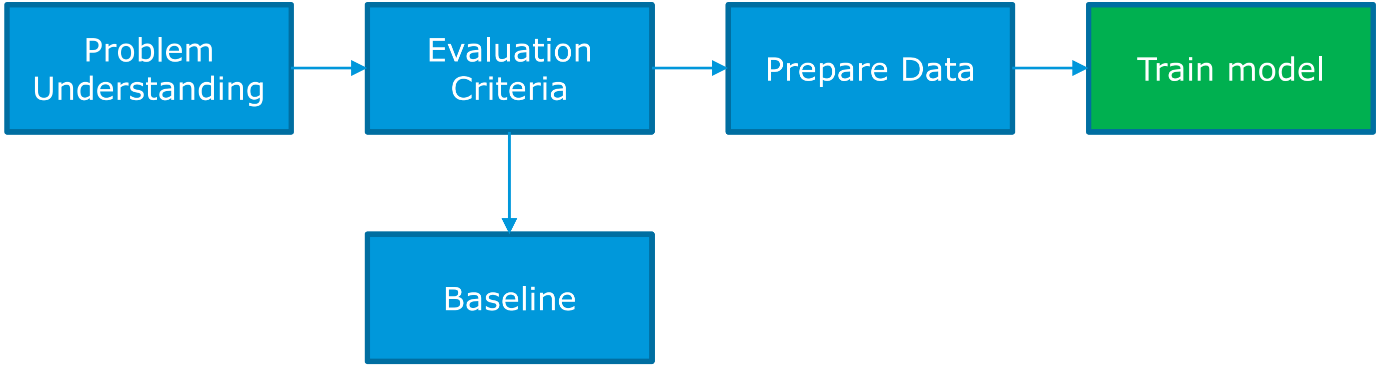
Machine Learning Workflow

- "Enough" data in the sense that it's representative of the behaviour the model needs to learn
- Features: data transformations that encode your knowledge



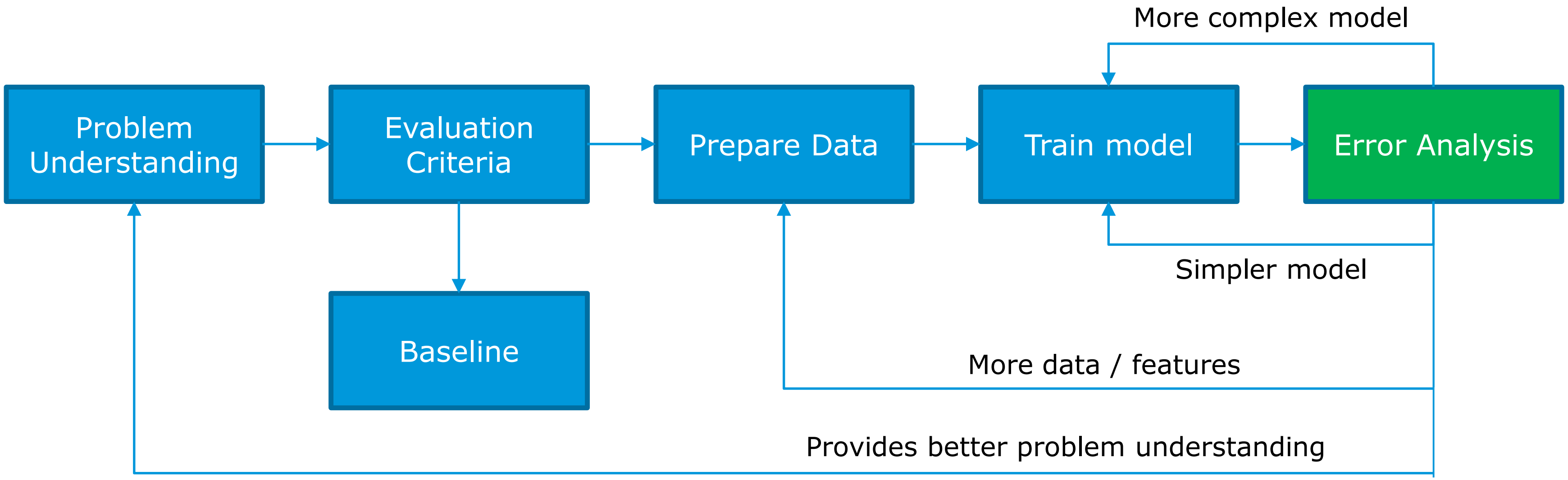
Machine Learning Workflow

- Use data / features to tune the parameters that optimize the evaluation criteria (e.g. minimise error)



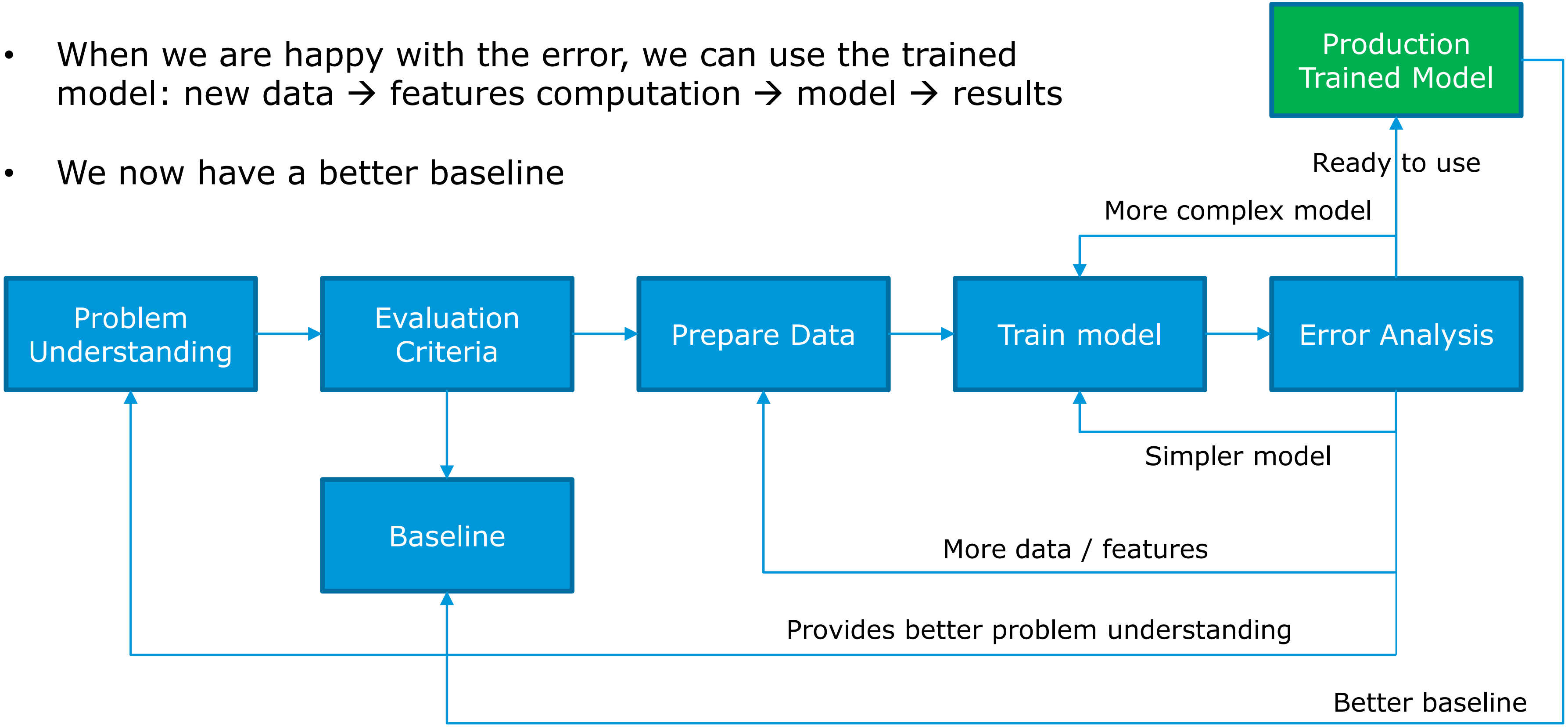
Machine Learning Workflow

- Understand what the model is doing: where is it right / wrong

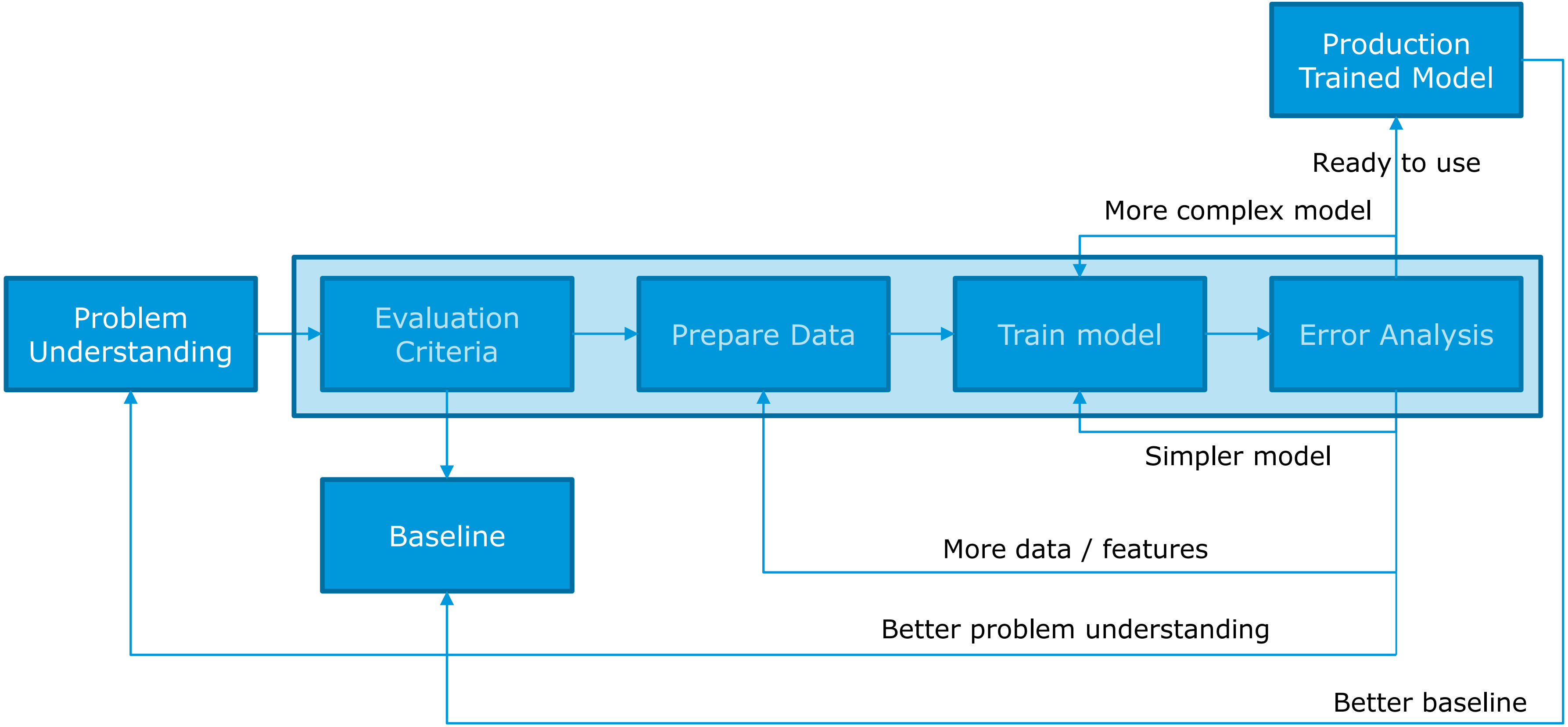


Machine Learning Workflow

- When we are happy with the error, we can use the trained model: new data → features computation → model → results
- We now have a better baseline



Machine Learning Workflow



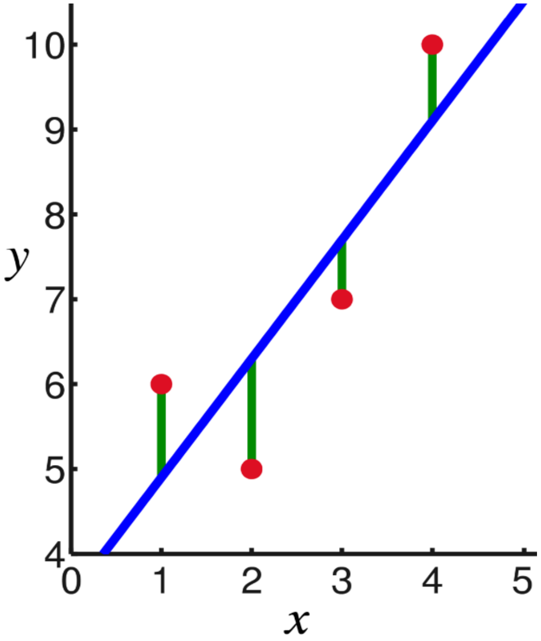
Evaluation Criteria

Typically a way to measure error.

Example:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2$$

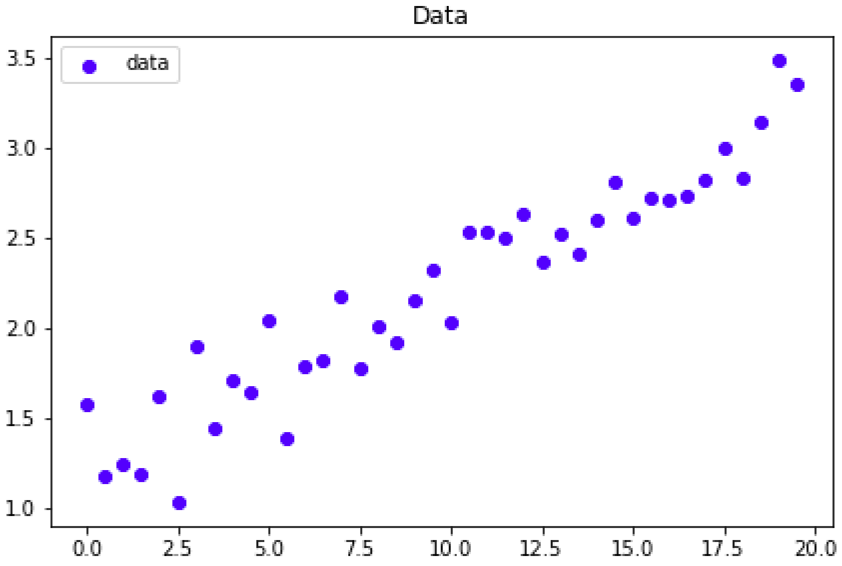
Mean Squared Error



* We will see more evaluation criteria in other sessions

Train Machine Learning Models

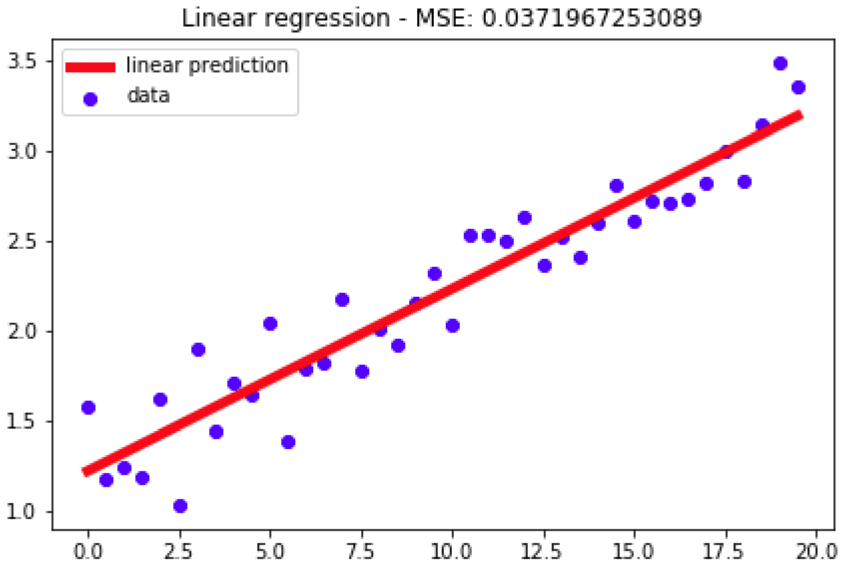
Let’s train 2 models to learn from this data



$$y = 0.1 * x + 1.25 + 0.2 * GaussianNoise$$

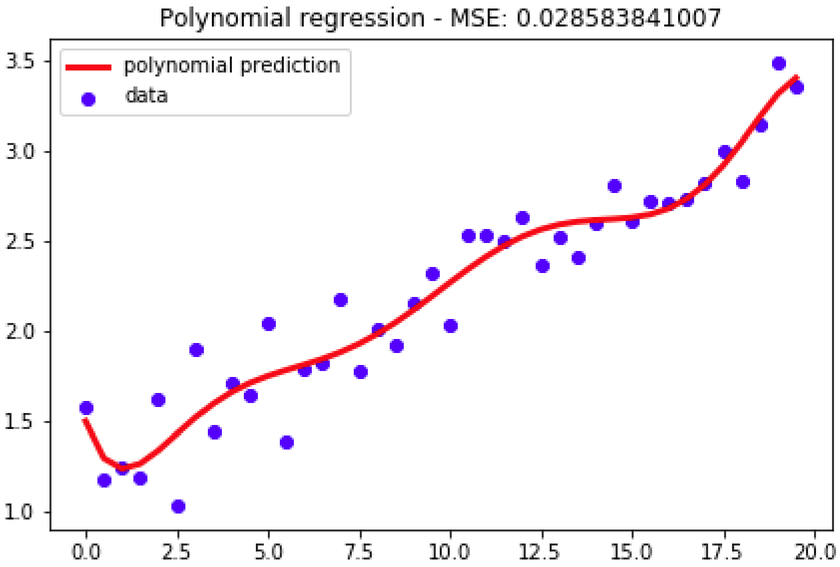
Train Machine Learning Models

Linear Regression



Mean Squared Error: 0.0372

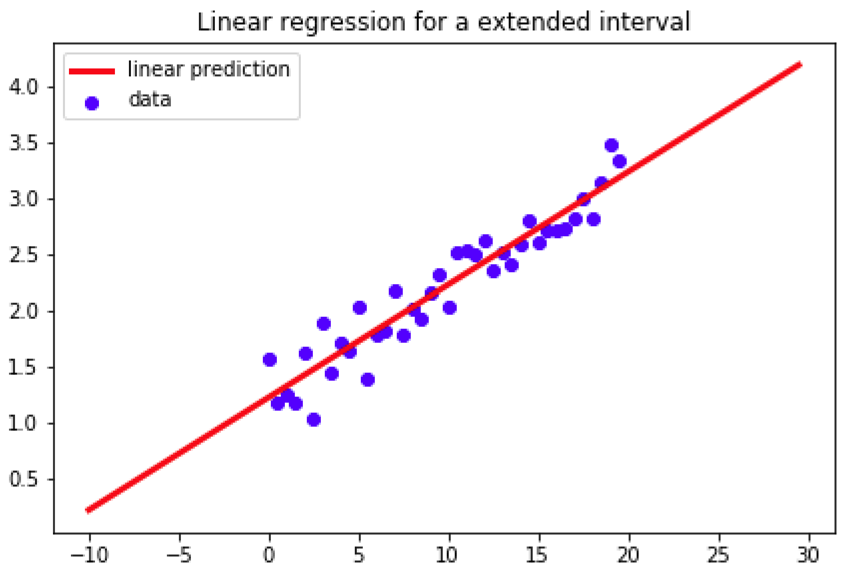
Polynomial Regression (degree 7)

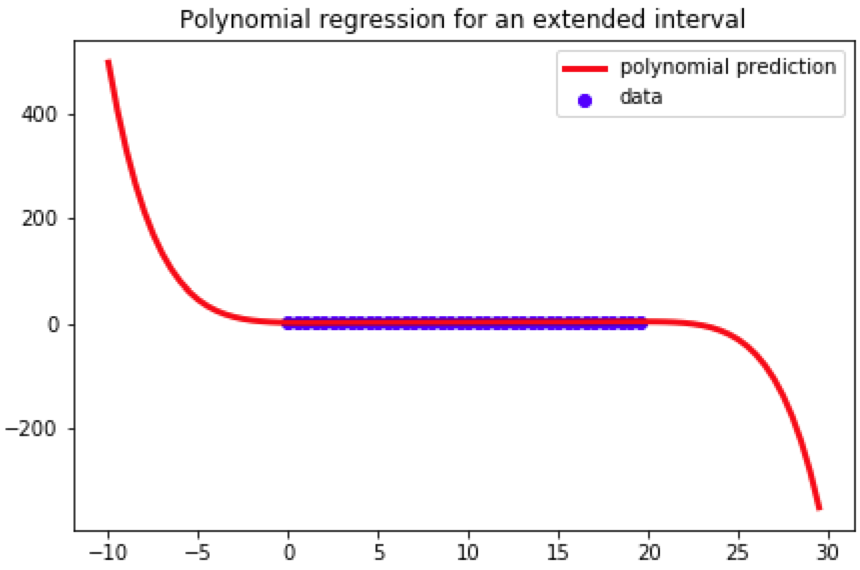


Mean Squared Error: 0.0286

Generalization

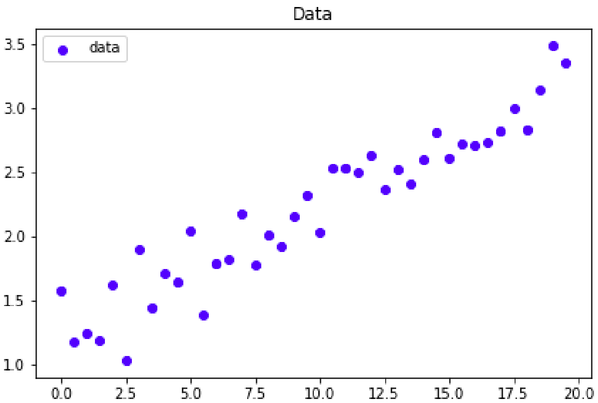
Generalization: the ability of a model to perform well on new data



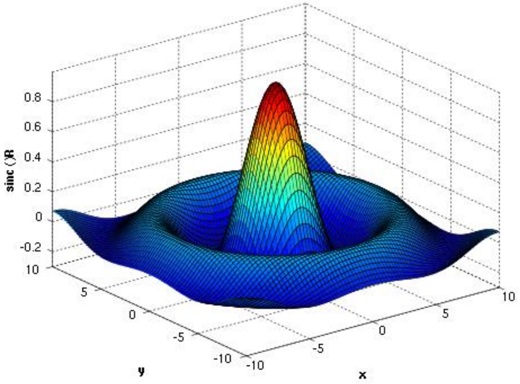


Generalization

How to tell if a Machine Learning model is generalizing well when we cannot visualize what it is doing?



1-dimensional data



2-dimensional data

?

n-dimensional data

Generalization

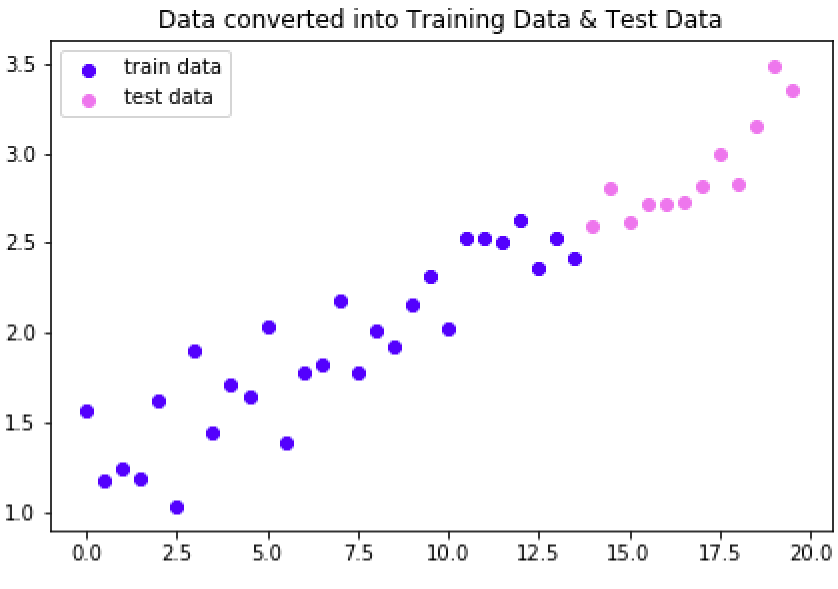
Split data in 2 sets

- Training Set: for training the model
- Testing Set: for evaluating the generalization of the model

* Validation set and cross-validation will be explained later

Generalization

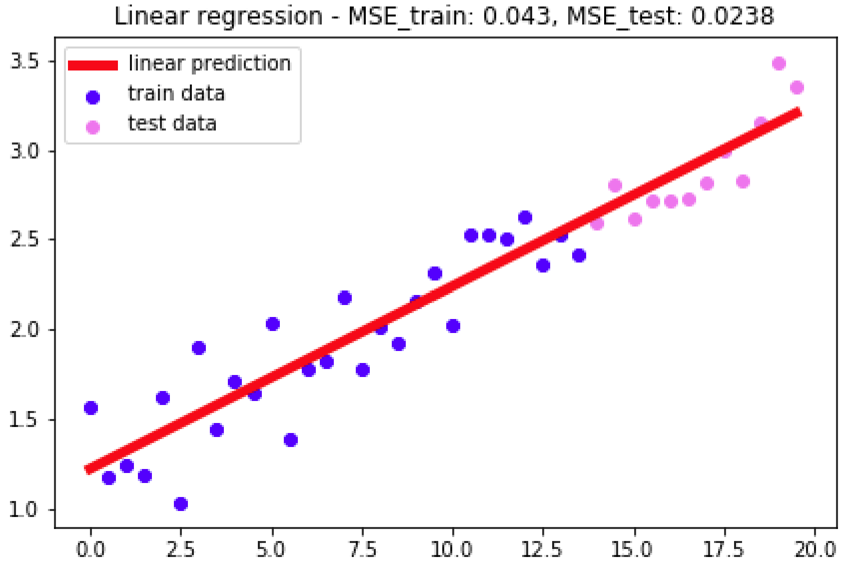
Data → Training Data + Testing Data



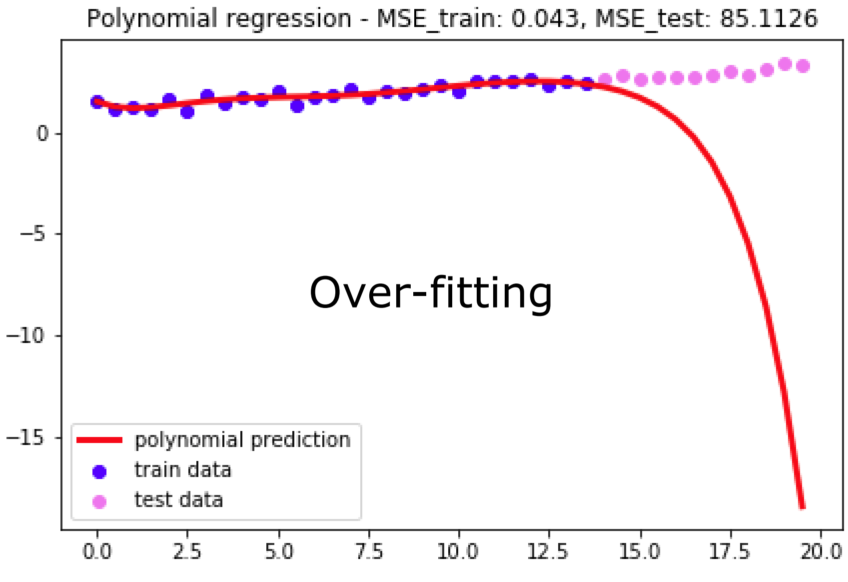
The Machine Learning model now learns **<u>from training data only</u>**

Generalization – Error Analysis

Evaluate Performance on Training Data & Test Data



MSE Train: 0.043, MSE Test: 0.0238



MSE Train: 0.043, MSE Test: 85.113

Generalization – Error Analysis - Guidelines

Train Error	Test Error	What to do?
Low Over-fitting	High	• Need a simpler model • Need more data (more data samples)

* High / Low compared with the error you are willing to accept

Generalization – Error Analysis - Guidelines

Train Error	Test Error	What to do?
Low	High	• Need a simpler model • Need more data (more data samples)
Over-fitting		

High	High	• Need more data
		  • Difficult to learn f(x, z) with only x. Get also z
		  • Get more data samples
		  • Additional features (e.g. $\frac{1}{x^2}$)
		• Need a more complex model

Under-fitting		

* High / Low compared with the error you are willing to accept

Generalization – Error Analysis - Guidelines

Train Error	Test Error	What to do?
Low Over-fitting	High	• Need a simpler model • Need more data (more data samples)

High Under-fitting	High	• Need more data
		  • Difficult to learn f(x, z) with only x. Get also z
		  • Get more data samples
		  • Additional features (e.g. $\frac{1}{x^2}$)
		• Need a more complex model

High	Low	Unusual: it could mean that the test data is too similar to the train data. Get more test data.

* High / Low compared with the error you are willing to accept

Generalization – Error Analysis - Guidelines

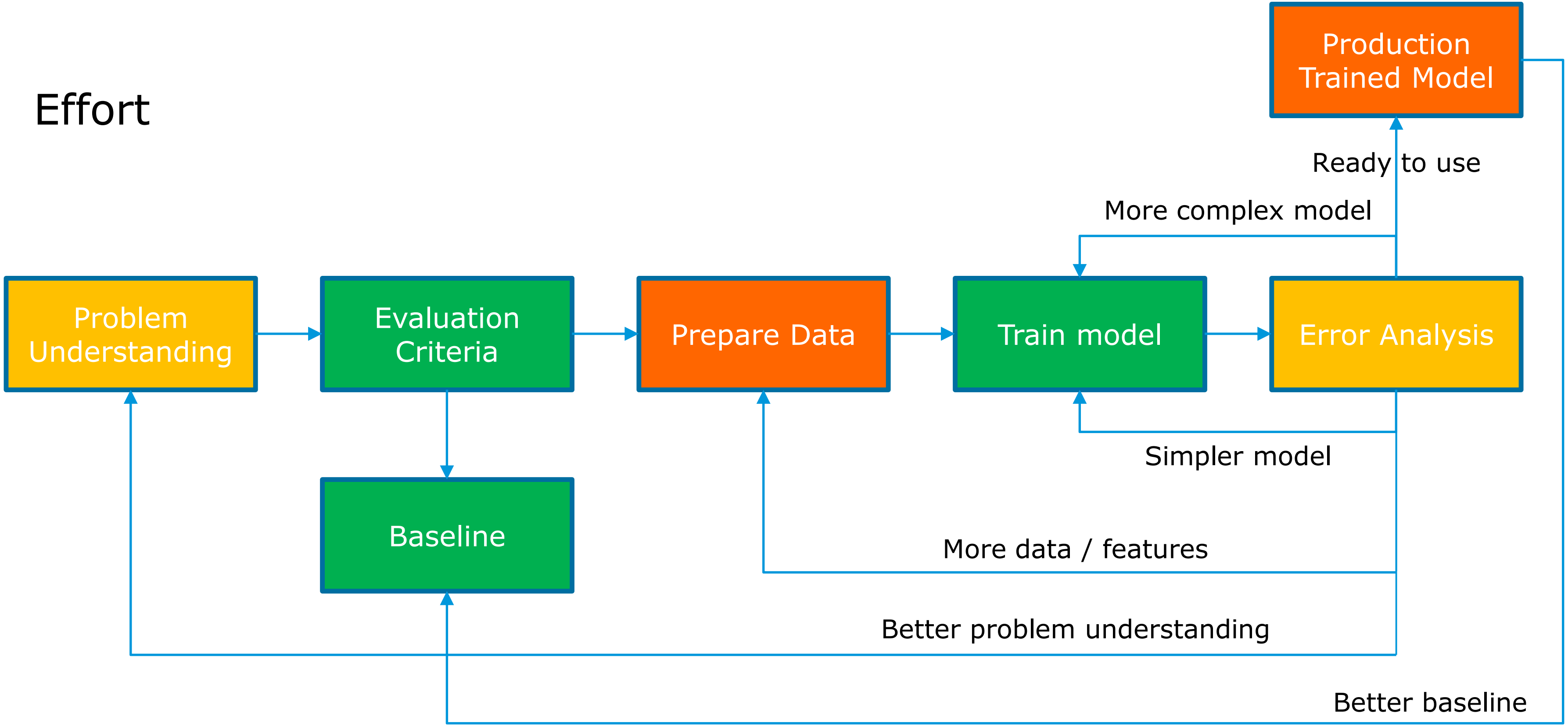
Train Error	Test Error	What to do?
Low	High	• Need a simpler model • Need more data (more data samples)
Over-fitting		

High	High	• Need more data
		  • Difficult to learn f(x, z) with only x. Get also z
		  • Get more data samples
		  • Additional features (e.g. $\frac{1}{x^2}$)
		• Need a more complex model

Under-fitting		
High	Low	Unusual: it could mean that the test data is too similar to the train data. Get more test data.
Low	Low	You're done! – congratulations

* High / Low compared with the error you are willing to accept

Machine Learning Workflow



Let’s recap

Today we have discussed about:

- What is machine learning
- Some applications in industry & space
- What is the ‘Learning’ in Machine Learning
- Types of machine learning
- Regression / Classification
- Machine Learning workflow
- Generalization

Lectures and Hands-on will be available on the Data Analytics ESA connect community

https://connect.esa.int/communities/community/data-analytics

What is next?

March 8th 16:00 – Press Room

Session 2: Supervised Learning (1)

- Linear, polynomial regression
- Lasso, Ridge regression
- Logistic Regression
- Support Vector Machines (SVM)

- Supervised Learning (1) in Python

Resources

Watch the video of this lecture

https://dlmultimedia.esa.int/download/public/videos/2048/03/001/4803_001_AR_EN.mp4

Get presentation and additional resources on

https://github.com/jmartinezheras/2018-MachineLearning-Lectures-ESA

Thank you

Data Analytics Team for Operations (DATO)

Jose Martinez Heras

LinkedIn: https://www.linkedin.com/in/josemartinezheras/