Accelerate Your Automation, Data & AI Journey!

#1: Build your first Machine Learning model and get started with Data Science

Speakers: Felix Augenstein, Data Scientist, IBM

Agenda:
1. Data Science
2. Artificial Intelligence
3. Machine Learning
4. ML Algorithms
5. ML Metrics
6. AutoAI vs Jupyter Notebook
7. Hands-on lab
8. Develop your own app with us
9. Q&A

IBM Developer

Data Science

Data Science

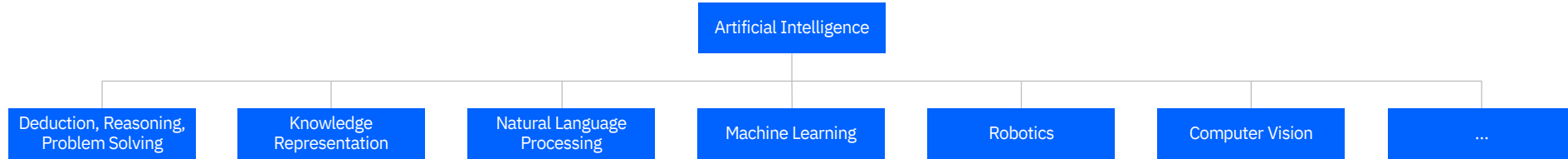Interdisciplinary field to extract knowledge and insights from structured and unstructured data.

(Source: Wikipedia)

CRoss **I**ndustry **S**tandard **P**rocess for **D**ata **M**ining (CRISP-DM):

Artificial Intelligence

Artificial Intelligence

```
                                    ┌─────────────────────────┐
                                    │ Artificial Intelligence │
                                    └─────────────────────────┘
```

Deduction, Reasoning, Problem Solving	Knowledge Representation	Natural Language Processing	Machine Learning	Robotics	Computer Vision	...

Artificial Intelligence (AI):
leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind.
(Source: https://www.ibm.com/cloud/learn/what-is-artificial-intelligence)
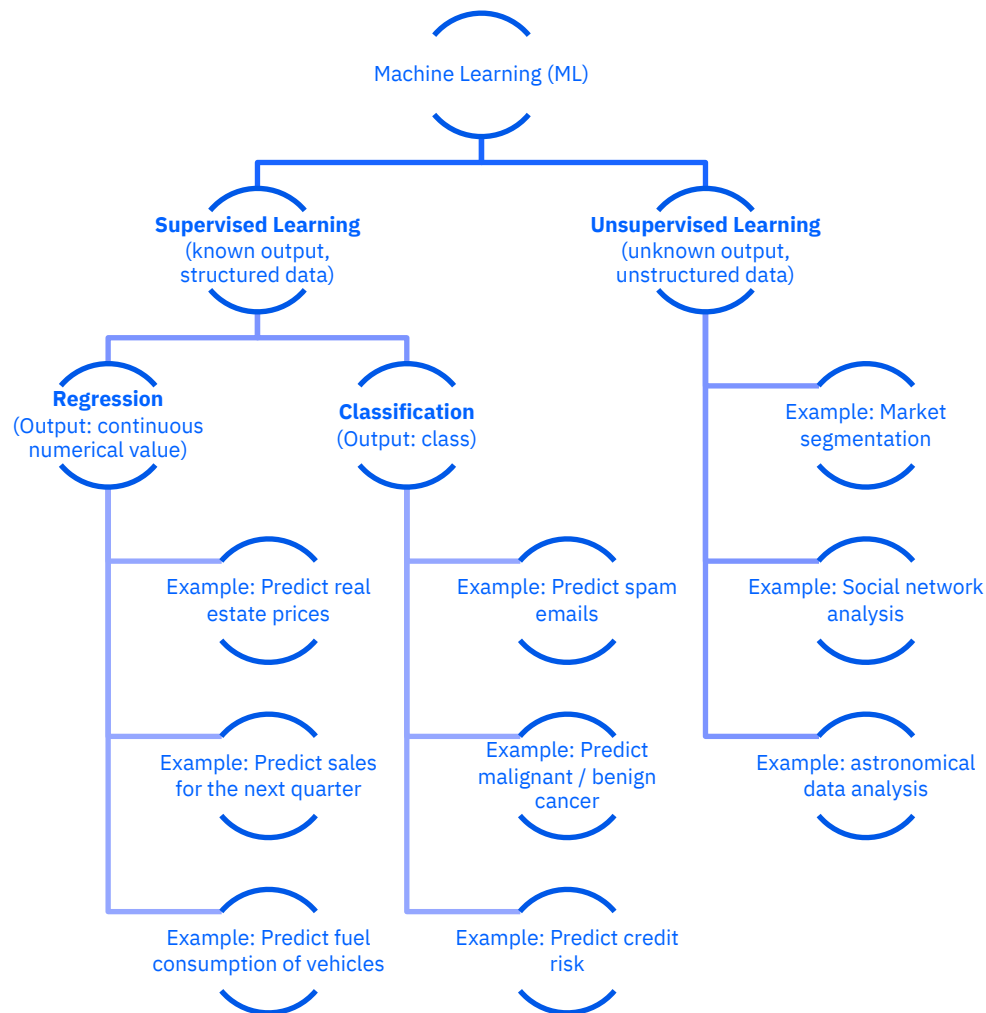
... or with other words:
to imitate human capabilities with computers

Machine Learning

Machine Learning

Field of study that gives computers the ability to learn without being explicitly programmed.

(Source: Arthur Samuel, 1959)

Machine Learning (ML)

Supervised Learning
(known output, structured data)

Unsupervised Learning
(unknown output, unstructured data)

Regression
(Output: continuous numerical value)

Classification
(Output: class)

Example: Market segmentation

Example: Predict real estate prices

Example: Predict spam emails

Example: Social network analysis

Example: Predict sales for the next quarter

Example: Predict malignant / benign cancer

Example: astronomical data analysis

Example: Predict fuel consumption of vehicles

Example: Predict credit risk

ML Algorithms

ML Algorithms

Linear Regression

Logistic Regression

Decision Trees

Support Vector Machines

Clustering

Naive Bayes

K-Nearest Neighbors

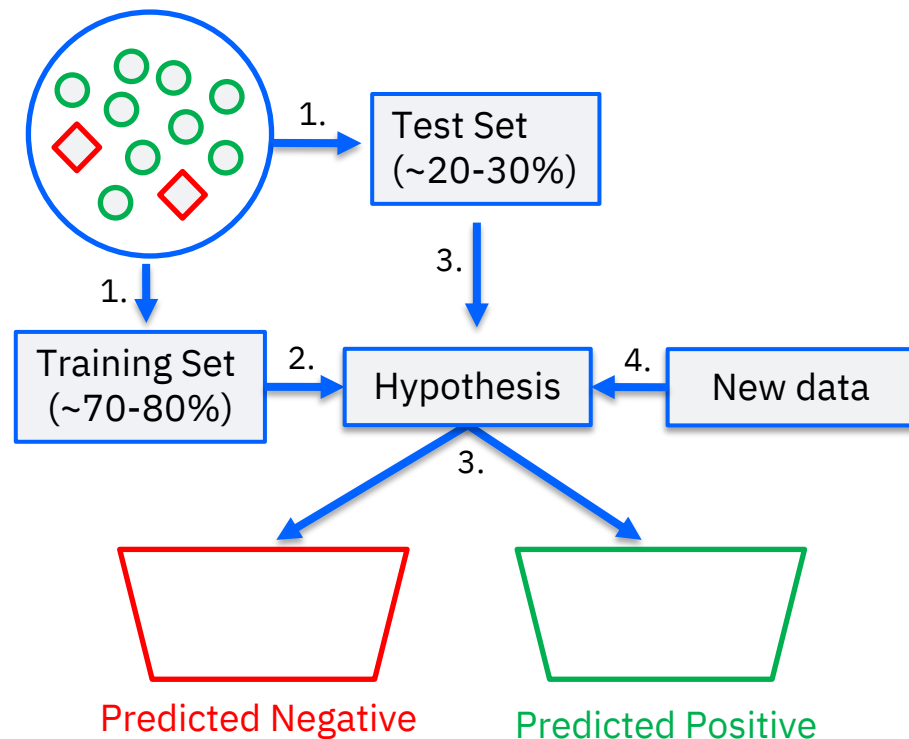Artificial Neural Networks

Random Forests

Boosting

And many more…

ML Metrics

ML Metrics

Accuracy, Recall, Precision & F1-Score for classification

Data Set (100%)

Test Set (~20-30%)

1.

Training Set (~70-80%)

2.

Hypothesis

3.

4. New data

3.

Predicted Negative

Predicted Positive

3. Evaluation

		Observed	
		Positive	Negative
Predicted	Positive	True Positive (tp)	False Positive (fp)
	Negative	False Negative (fn)	True Negative (tn)

Precision

Recall

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

$$\text{Recall} = \frac{tp}{tp + fn}$$

$$\text{Precision} = \frac{tp}{tp + fp}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

ML Metrics

RMSE for regression

Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit. The lower or closer the RMSE is to 0, the better the model or algorithm.

$$RMSE = \sqrt{\overline{(f - o)^2}}$$

Where:

- f = forecasts (expected values or unknown results),
- o = observed values (known results).

Other metrics include Mean Square Error (MSE), Mean Absolute Error (MAE), Median Absolute Error (MedAE), etc.
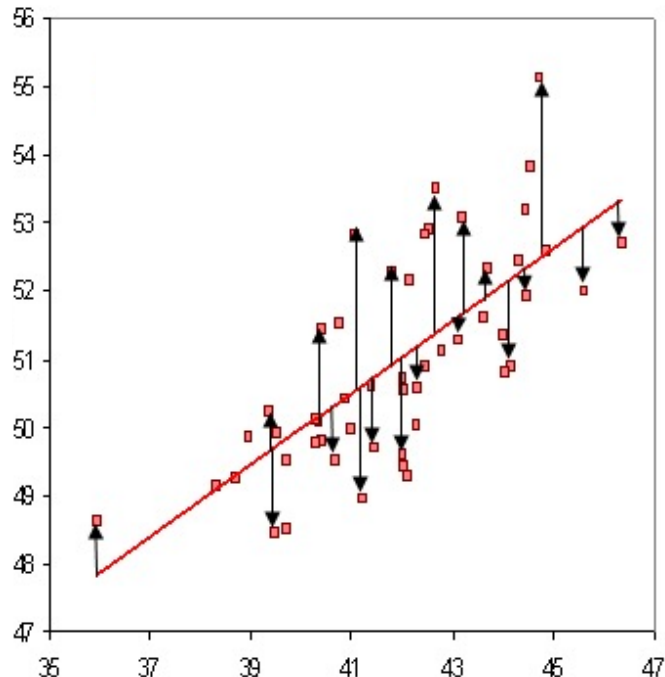
AutoAI vs Jupyter Notebooks

AutoAI vs Jupyter Notebook

IBM's Hybrid Cloud & AI solutions

Business Transformation and Hybrid Cloud Services	IBM Consulting

Hybrid Cloud Software	IBM Software — IBM Cloud Paks — Cloud Pak for Data (as a Service) / ... — Watson Machine Learning / **Watson Studio** / Watson OpenScale / ... — **AutoAI** / **Jupyter Notebooks** / ...

Hybrid Cloud Platform	Red Hat Hybrid Cloud Platform — Red Hat OpenShift

Infrastructure	IBM Cloud	Public Clouds AWS | Azure | Other	IBM Systems	Enterprise Infrastructure

AutoAI vs Jupyter Notebook
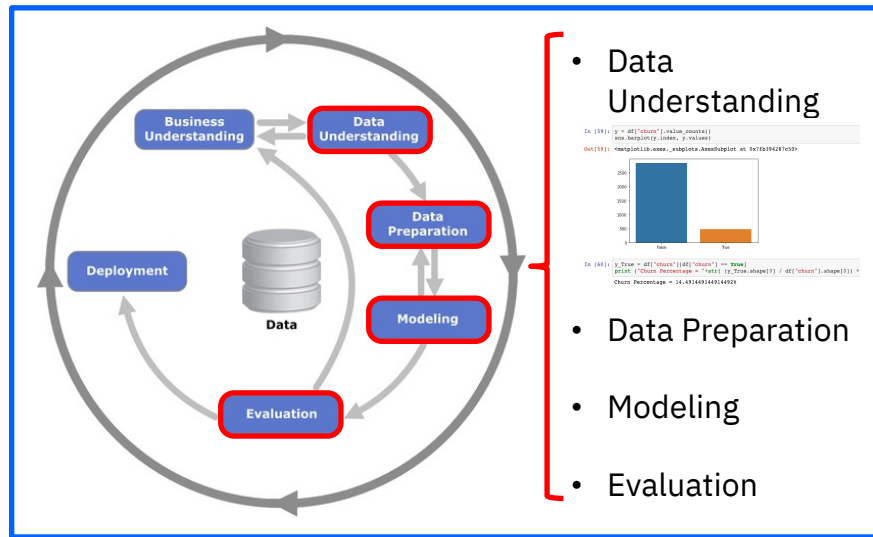
Example: Predict Customer Churn (Supervised Learning → Classification)
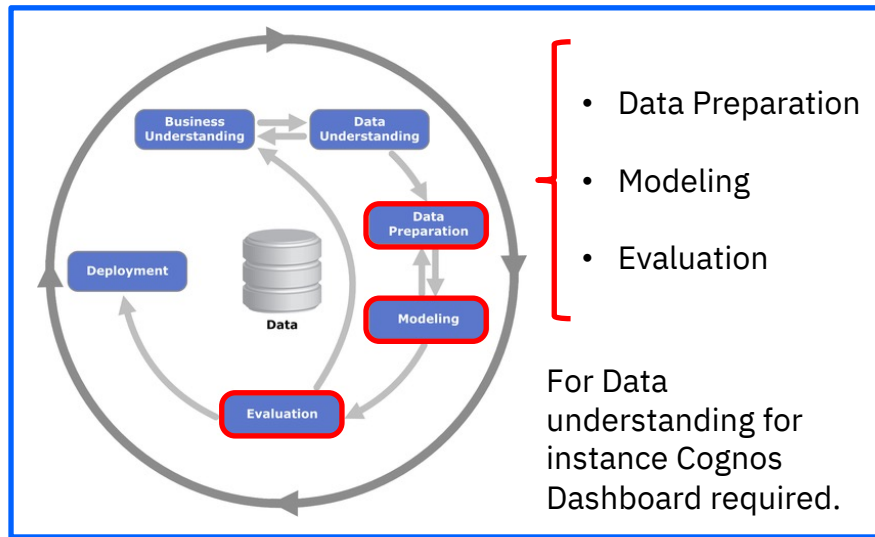
AutoAI

Jupyter Notebook

AutoAI vs Jupyter Notebook

Example: Predict Customer Churn (Supervised Learning → Classification)

AutoAI

- Data Preparation
- Modeling
- Evaluation

For Data understanding for instance Cognos Dashboard required.

Jupyter Notebook

- Data Understanding
- Data Preparation
- Modeling
- Evaluation

AutoAI vs Jupyter Notebook

Example: Predict Customer Churn (Supervised Learning → Classification)

AutoAI

+ Fast & Easy

+ Suitable for technical- & business experts

- Not 100% customizable

- No visualizations of data

Jupyter Notebook

+ 100% Customizable

+ Visualizations of data

- Difficult and time consuming

- Only for technical experts

Hands-on Lab

Hands-on Lab

Example: Predict Customer Churn (Supervised Learning → Classification)

Option 1: AutoAI

Develop a ML model using AutoAI, you can use a Cognos Dashboard to visualize the data.

Tutorial I + II

Option 2: Jupyter Notebook

Develop a ML model using Jupyter Notebooks.

Tutorial I + IV

https://github.com/FelixAugenstein/cloud-pak-for-data-tutorial

Develop your own app with us

Develop
your
own app with us!

Now it is your turn, send us your message via Email: **OR** **via LinkedIn:**

Mailto: MKRAUSE@de.ibm.com, Felix.Augenstein@ibm.com

Subject: Develop your own app - <YOUR_COMPANY/>

As a direct message to Marion or Felix

Hi Marion / Hi Felix,

I participated in the Accelerate your Automation, Data & AI Journey!

I would like to reserve an individual session for my 10 - 15-minute pitch for the 22. or 23. of February 2022.

My use case is <YOUR_USE_CASE/> and of the technologies covered, I am interested in the following <TECHNOLOGY/>.

Regards,
<YOUR_NAME_AND_COMPANY/>

Marion (Krause) Nehring
IBM Hybrid Cloud Build Team & Developer Advocacy DACH | Client & Ecosystem Lead, Program Manager at IBM
linkedin.com/in/marionnehring/

Felix Augenstein
Data Scientist, Hybrid Cloud Build Team at IBM
linkedin.com/in/felixaugenstein/

Idea
time

Customers who bought this item also bought:

SPAM

5434
5434

Q&A