Basics of Machine Learning

20_KIN2 – Artificial Intelligence and Machine Learning

Lecture Contents

- Machine Learning & Learning
- Applications & Examples of ML
- Induction, Transduction, Deduction
- Supervised & Unsupervised Learning
- Experience

Terms

- Artificial Intelligence
- Machine Learning
- Optimization
- Simulation
- Modeling
- Robotics
- Data Science
- Data Mining
- Data Analytics
- Autonomous Decisions

Consequences

Apple Card algorithm sparks gender bias allegations against Goldman Sachs

Entrepreneur David Heinemeier Hansson says his credit limit was 20 times that of his wife, even though she has the higher credit score

By **Taylor Telford**

November 11, 2019 at 4:44 p.m. GMT+1

Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT

Via The Guardian | Source TayandYou (Twitter)

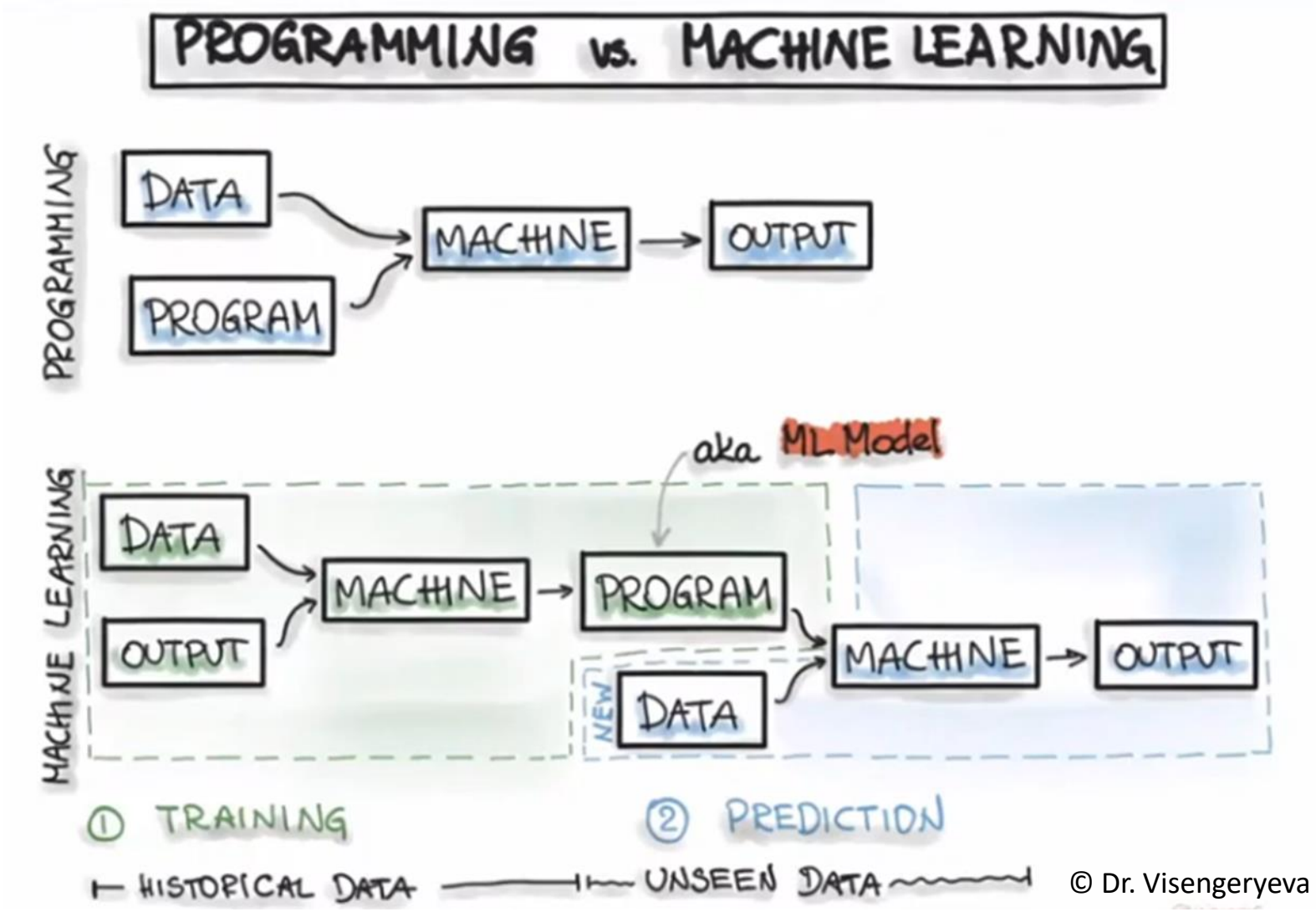
AMS-Algorithmus: Forscher warnen vor Diskriminierung und bemängeln fehlende Transparenz

Im Juli dürfte der Algorithmus, mit dem das Arbeitsmarktservice Jobsuchende in drei Gruppen einteilt, in Echtbetrieb gehen. Einige Wissenschafter äußern Bedenken

András Szigetvari 25. Februar 2020, 06:00 512 Postings

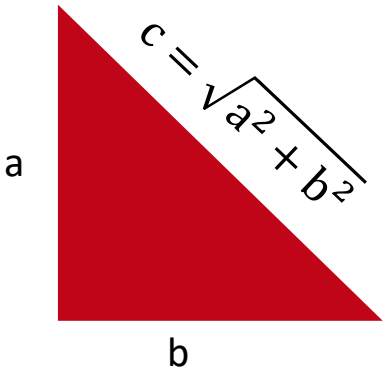
ProPublica

https://www.youtube.com/watch?v=VYlXNWxqJ2A



© Dr. Visengeryeva

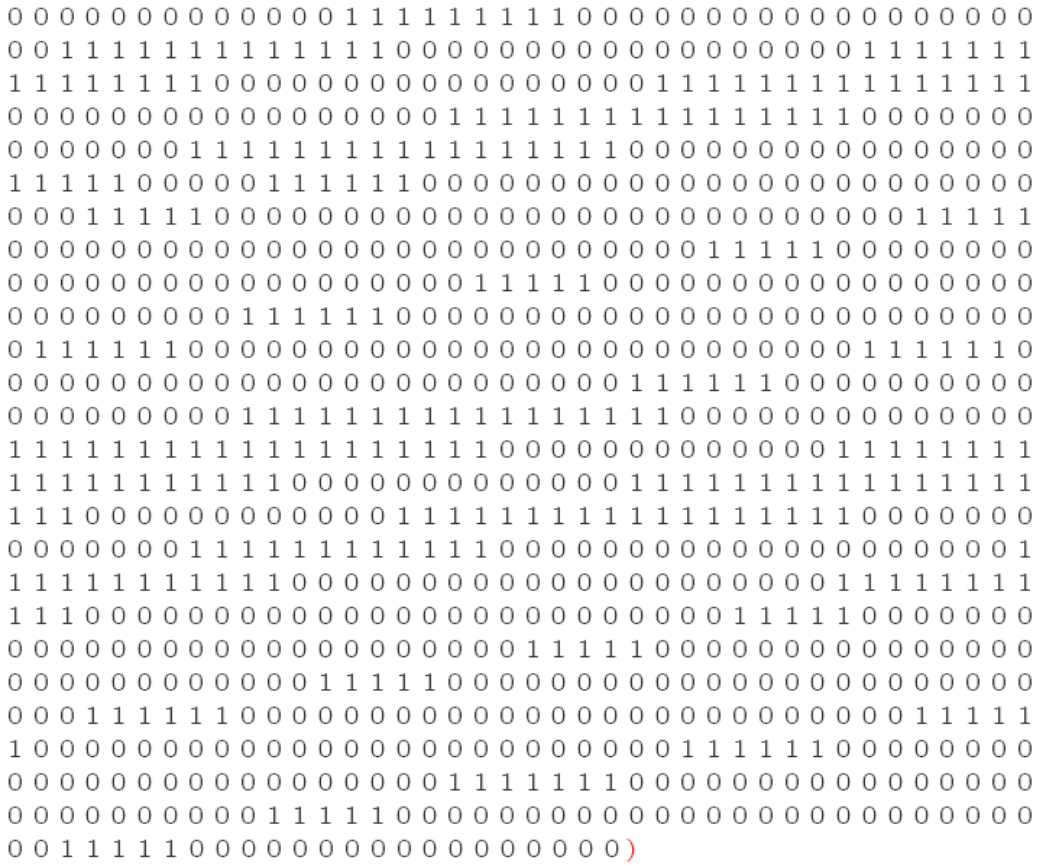
Machine Learning



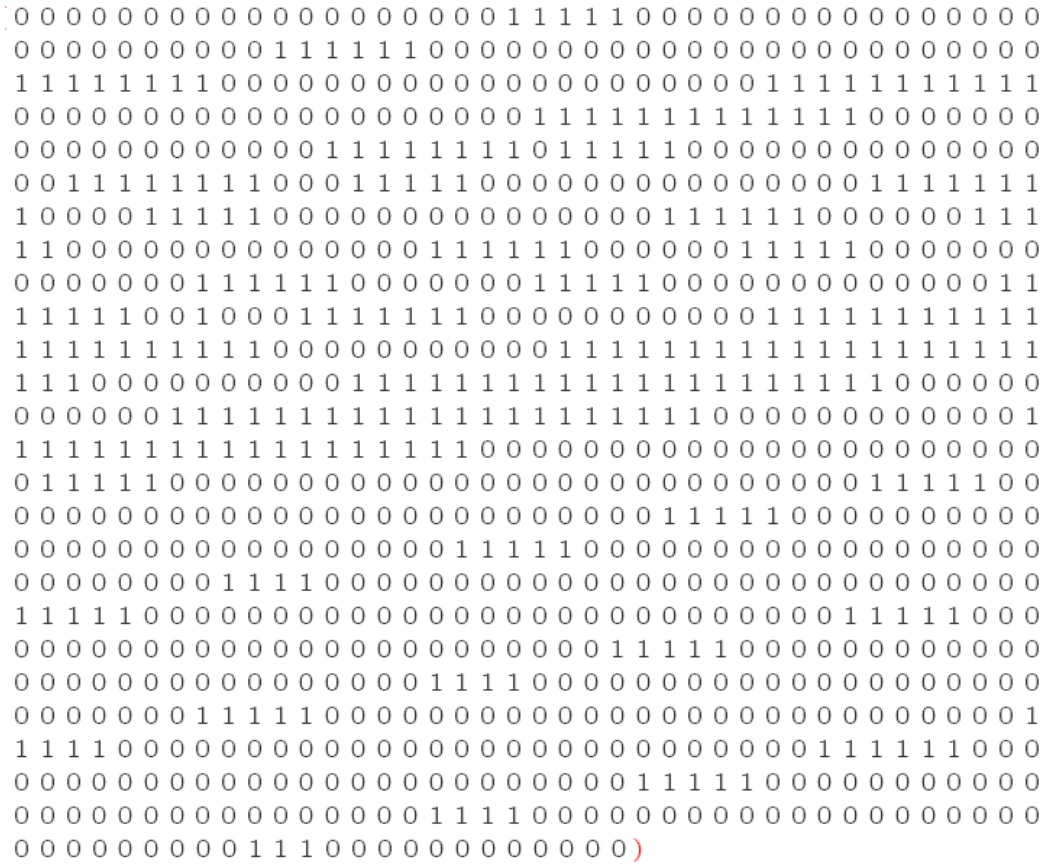
$$c = \sqrt{a^2 + b^2}$$

Number recognition

7

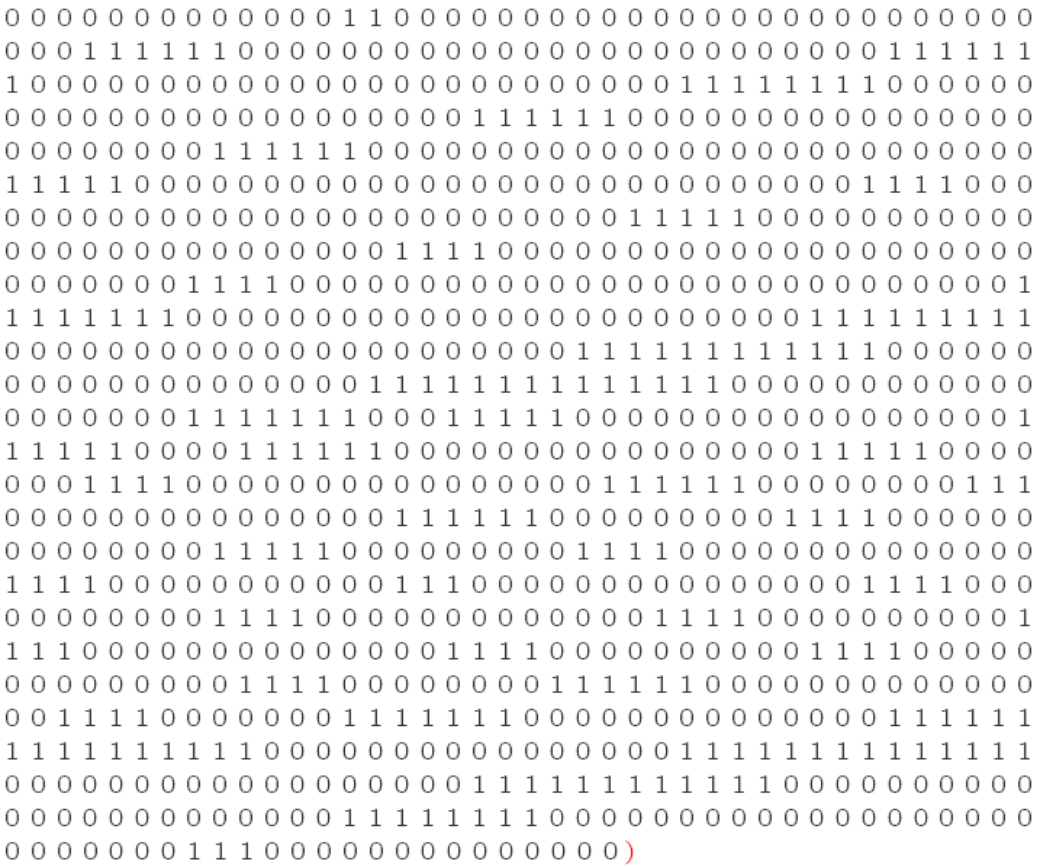


4

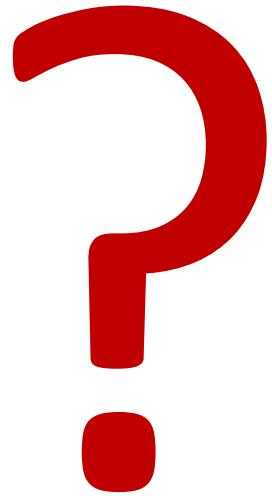


Prognosis

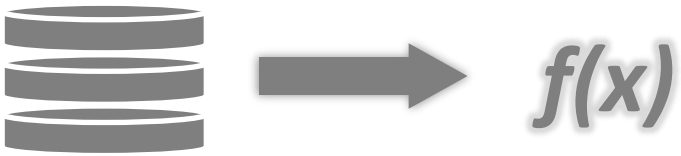
Input



Output



Machine Learning in Python




```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv("data.csv")
X = data.iloc[:, :-1]
y = data.iloc[:, -1]
```

```python
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X, y)

from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
model  = LinearDiscriminantAnalysis(store_covariance = True)
model.fit(X, y)

from sklearn.tree import DecisionTreeRegressor
model  = DecisionTreeRegressor(max_depth = 5)
model.fit(X, y)

from sklearn.ensemble import RandomForestClassifier
model  = RandomForestClassifier(n_estimators=100)
model.fit(X, y)

from sklearn.svm import SVR
model  = SVR(kernel = "linear")
model.fit(X, y)

predictions = model.predict(X)
```

Machine Learning / Data Mining

"The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience."

(Tom Mitchell, 1997)

"Data Mining is the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data."

(Fayyad *et al*, 1996)

What is learning?

"Learning is making useful changes in our minds."

(Marvin Minsky, 1985)

"Learning is constructing or modifying representations of what is being experienced."

(Ryszard Michalski, 1986)

"Learning means behaving better as a result of experience."

(Stuart Russel & Peter Norvig, 1985)

"Learning denotes changes in a system that are adaptive in the sense that they enable the system to do the same task or tasks drawn from the same population more efficiently next time."

(Herbert Simon, 1983)

Formalization of Learning

Supervised Learning

The system gets questions (Inputs) with correct answers (Outputs) with the goal that after the training the system can give correct answers to new questions.

Unsupervised Learning

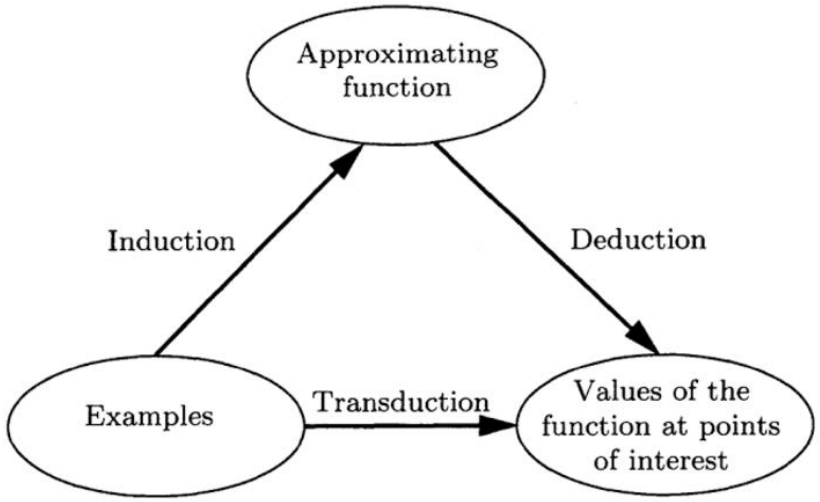
System has to identify the structure of data itself and identify patterns.

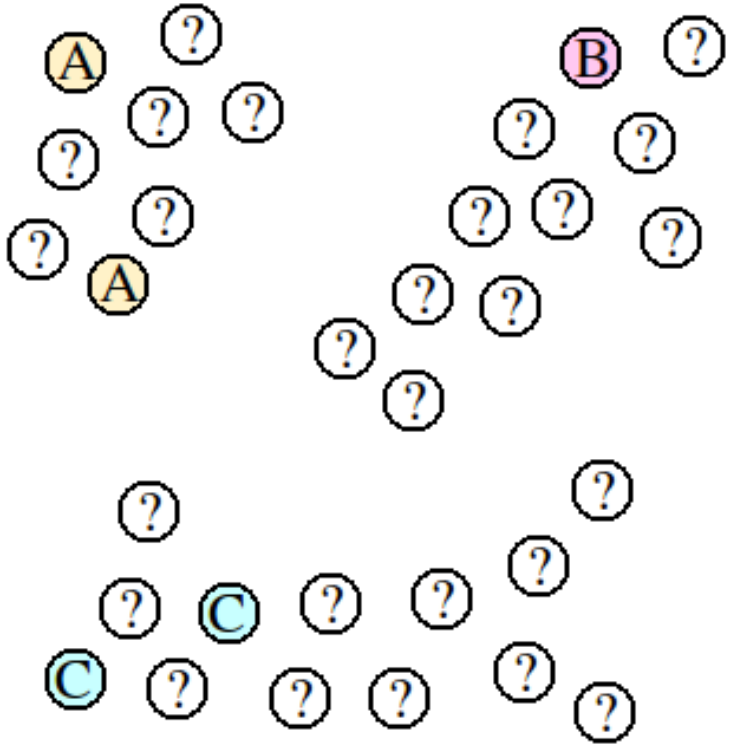
Reinforcement Learning

Control an agent by taking actions in an environment to maximize a specific reward.

Transduction

The process of directly drawing conclusions about new data from previous data, without constructing a model.

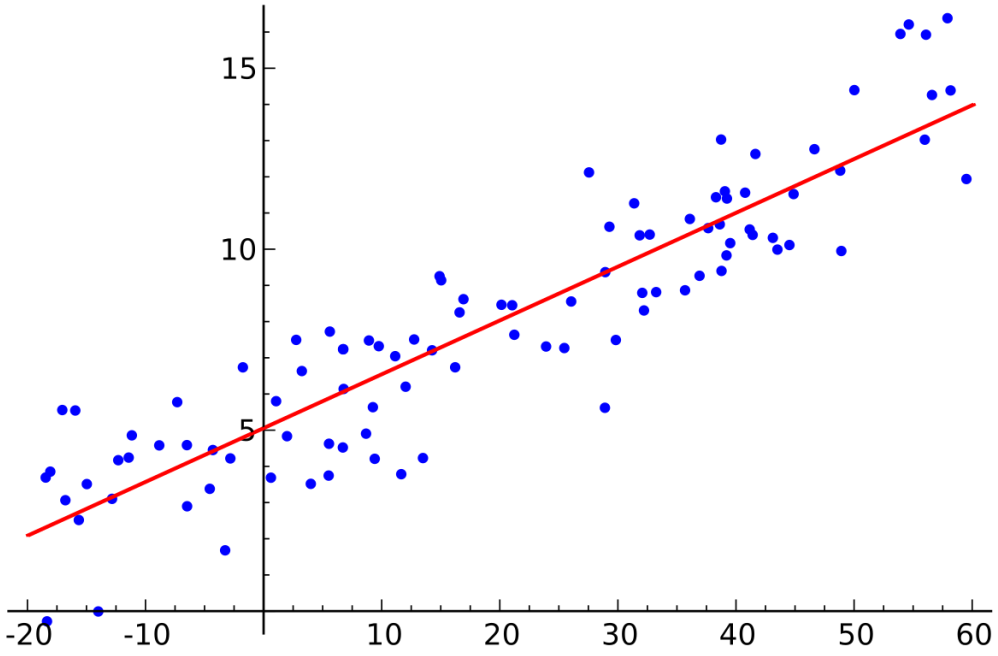


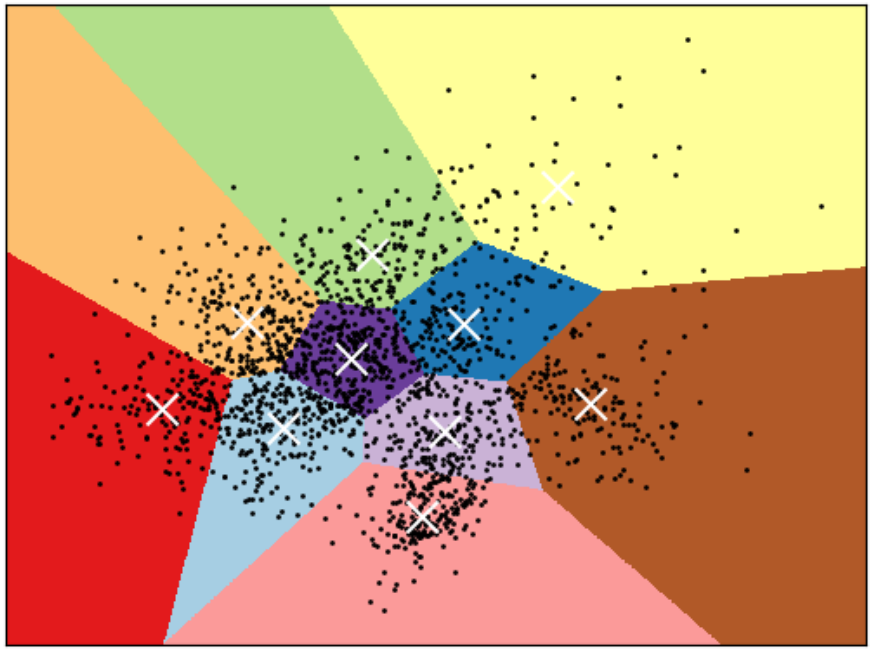


Induction

Induction = Inferring from specific cases to the general

Deduction = Applying knowledge to new cases





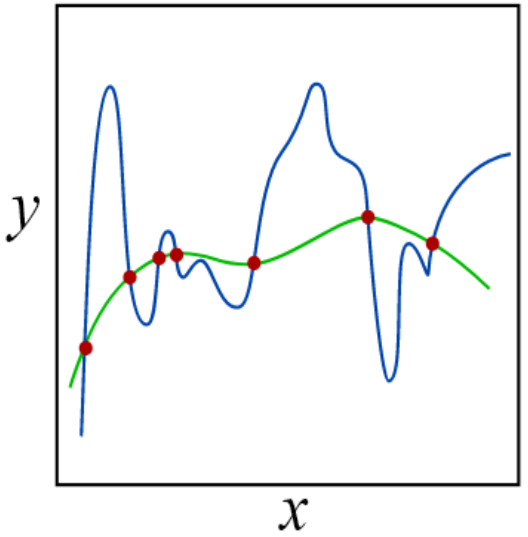
Supervised Learning

Given: A set of examples (x_i, y_i)

$S = \{(x_1, y_1), (x_2, y_2), \dots, (xn, yn)\} \subseteq x \times y$

Transduction: Output y for input x_0 ??

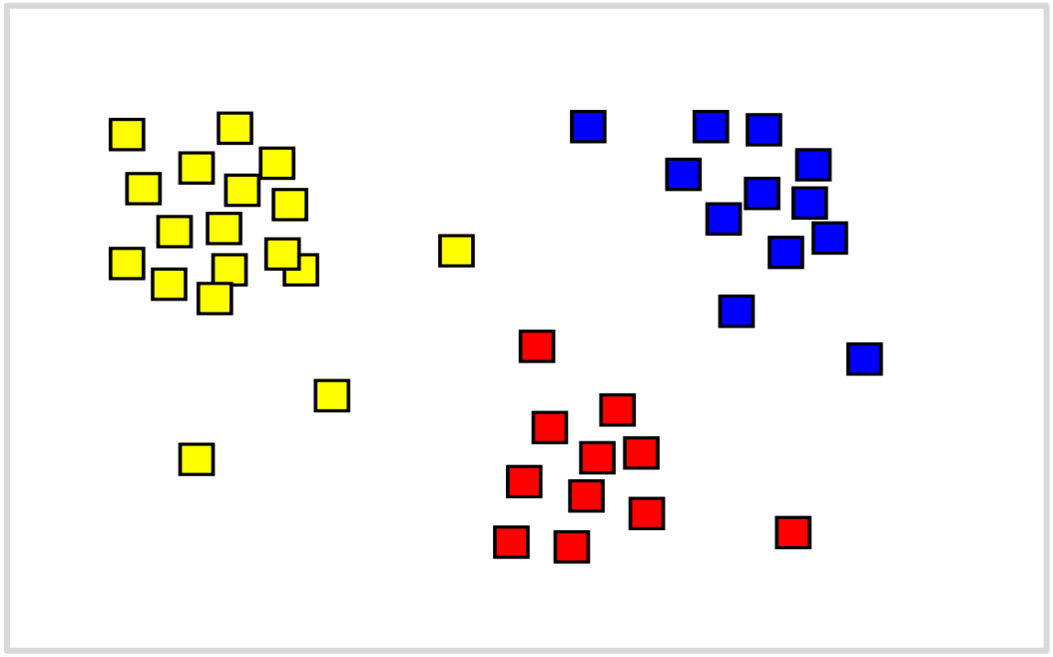
Induction: Complete functionally relation $f : X \rightarrow Y$??

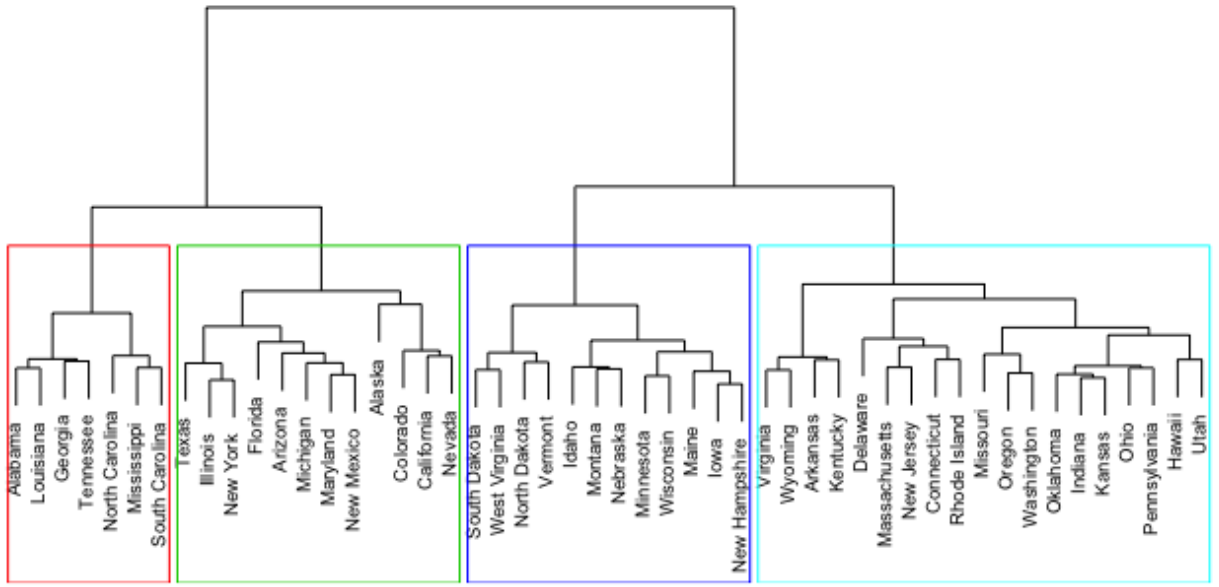


Unsupervised Learning

Given: A set of examples (X)

Output: Patterns / Commonalities in the data X

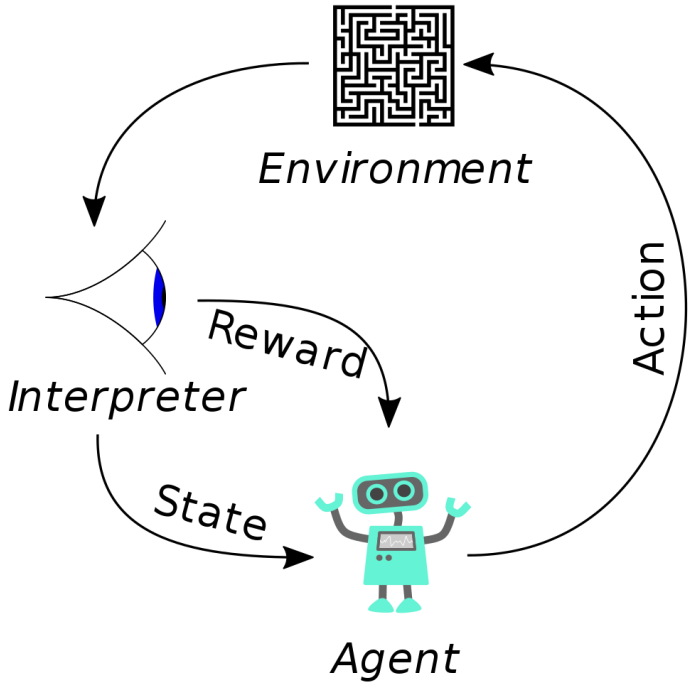


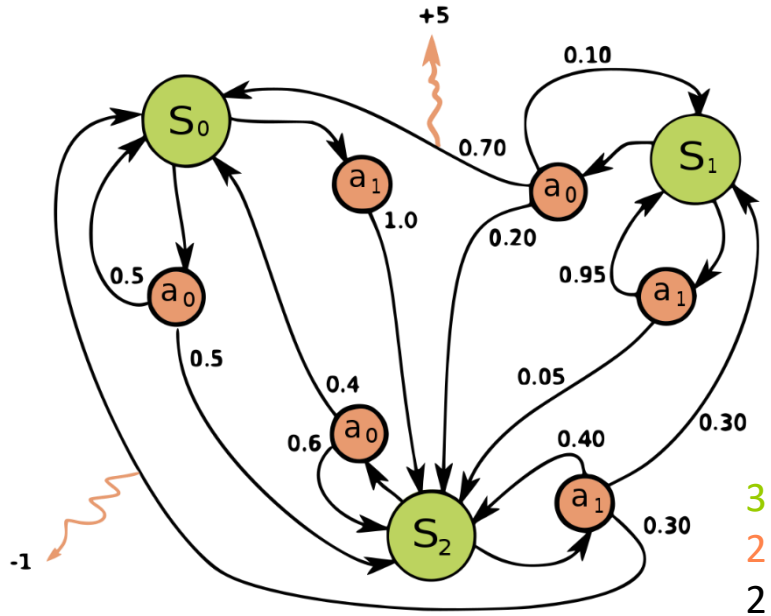


Reinforcement Learning

Given: Agent, Environment, Reward function

Output: Learnt behavior (policy)





3 States (S_i)

2 Actions (a_i)

2 Rewards $(+5, -1)$

Unsupervised machine learning tasks

Distinction based on distance metrics

Clustering

- Group together unlabeled data
- Similarity measure (metric) required

Dimensionality Reduction

- Obtain low-dimensional representation
- Retain meaningful information

Association

- Collocation extraction
- Find connections between terms

Supervised machine learning tasks

Distinction based on output

Classification

- Categorical output
- Binary classification
- Multi-class classification
- Multi-label classification

- Ordinal classification

Regression

- Numerical output
- Linear regression
- Nonlinear regression

- Time series analysis

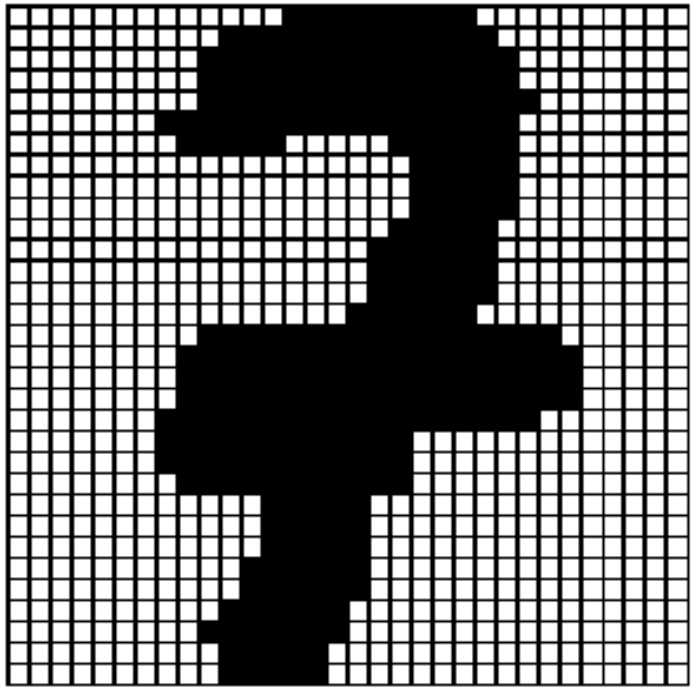
Examples of Machine Learning

List examples of machine learning applications that you encounter in your every day life

- Why is the example connected to machine learning?
- What is the method behind the application?
- What is necessary for the method to succeed?

ML Examples

Recognize handwritten numbers



System properties

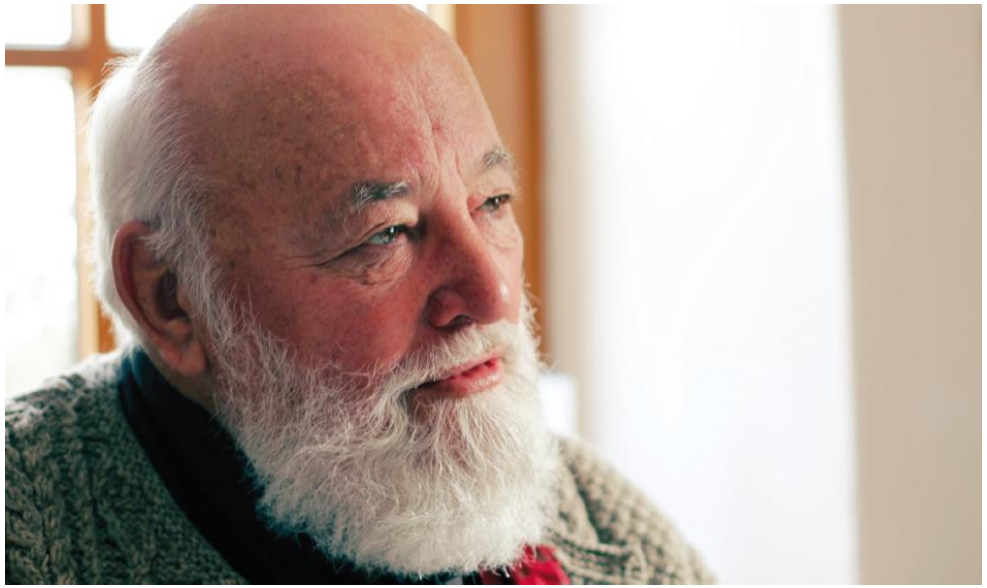
Input: $X \in \{0, 1\}^{1024}$

Output: $Y \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

F(X) ?

ML Examples

Classification of man or woman?



System properties

Input: Picture of a person

Output: Sex of the person

F(X) ?

ML Examples

Classification of man or woman

System properties

Input: Picture of a person

Output: Sex of the person

F(X) ?

ML Examples

Object recognition

System properties

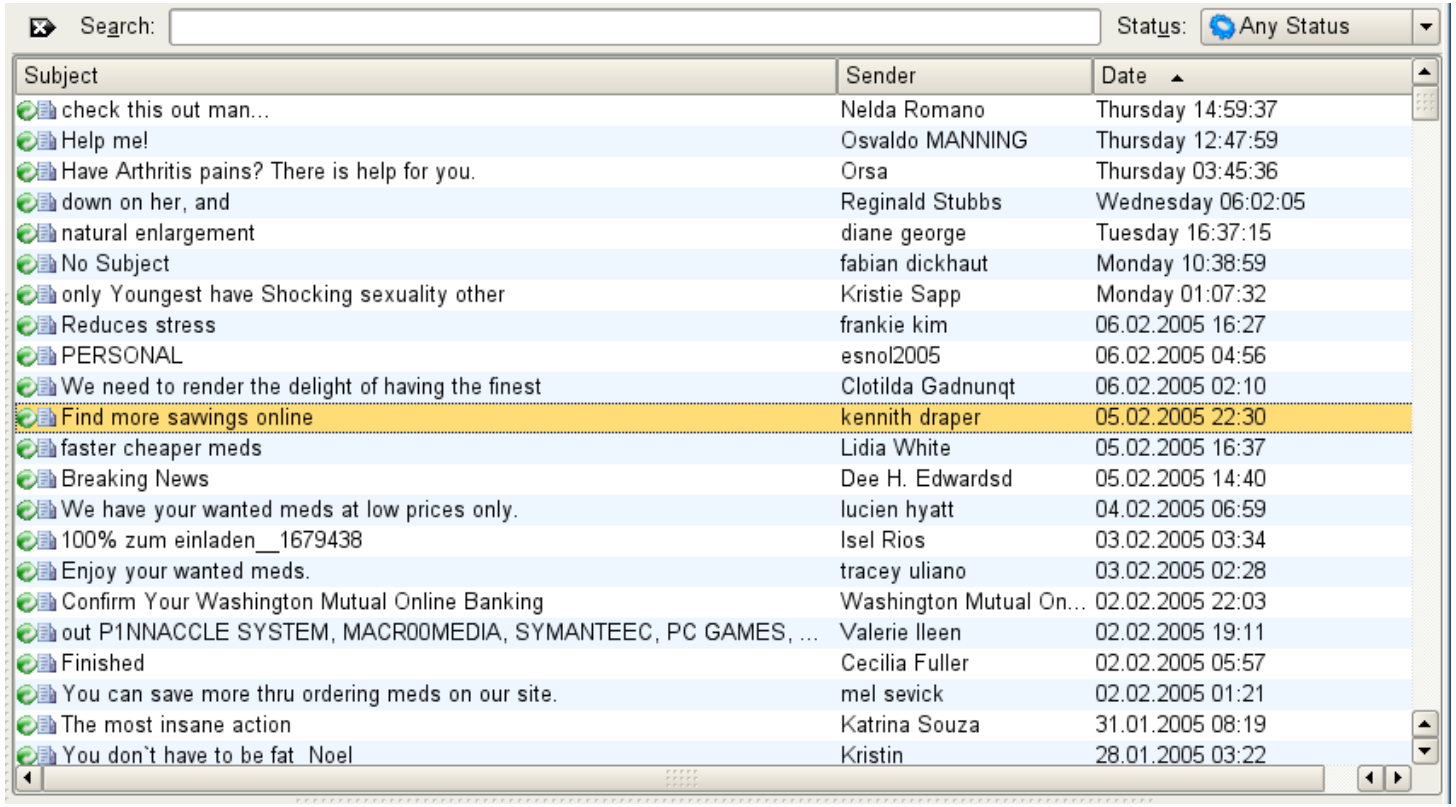
Input: Picture

Output: List of objects

F(X) ?

ML Examples

Spam detection



System Properties

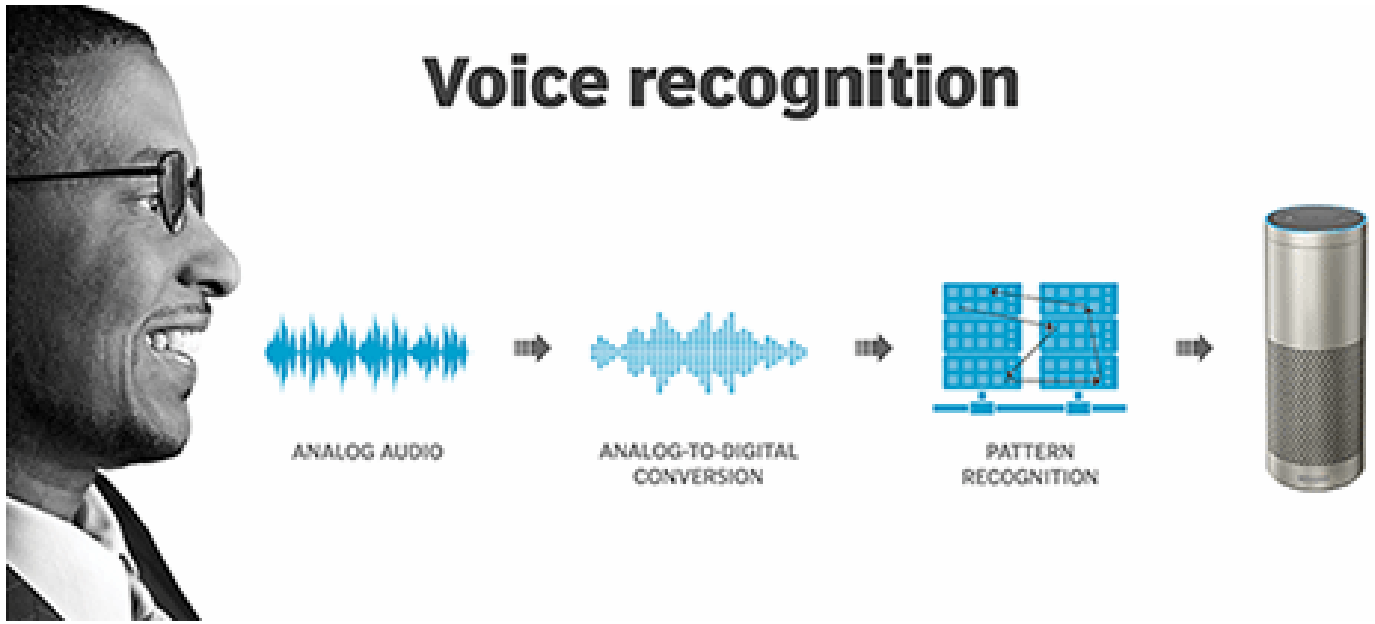
Input: E-Mails (+Metadata)

Output: Spam or No Spam

F(X) ?

ML Examples

Speech Recognition



System Properties

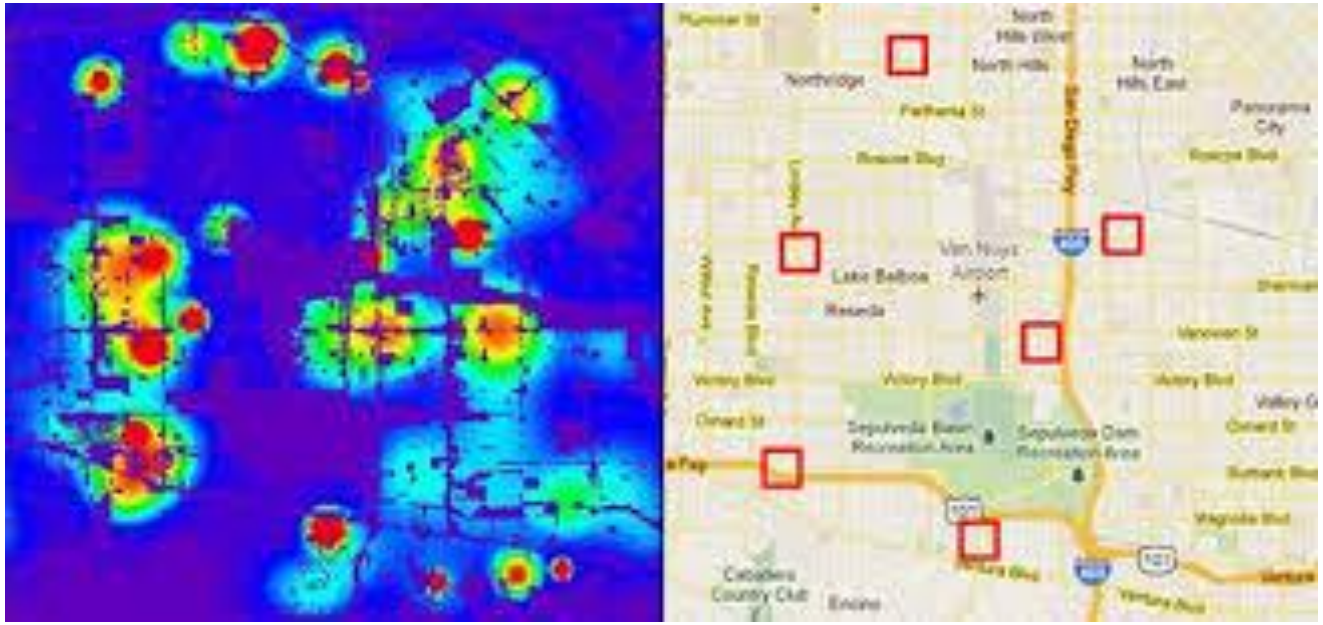
Input: Audio files

Output: Recognized text

F(X) ?

ML Examples

Predictive Policing



System Properties

Input: Lots of data sources

Output: Predictions from crime probability

F(X) ?

ML Examples

Self driving cars

System Properties

Input: Sensors, Maps, Route, …

Output: Driving actions

F(X) ?

ML Examples

Playing video games

System properties

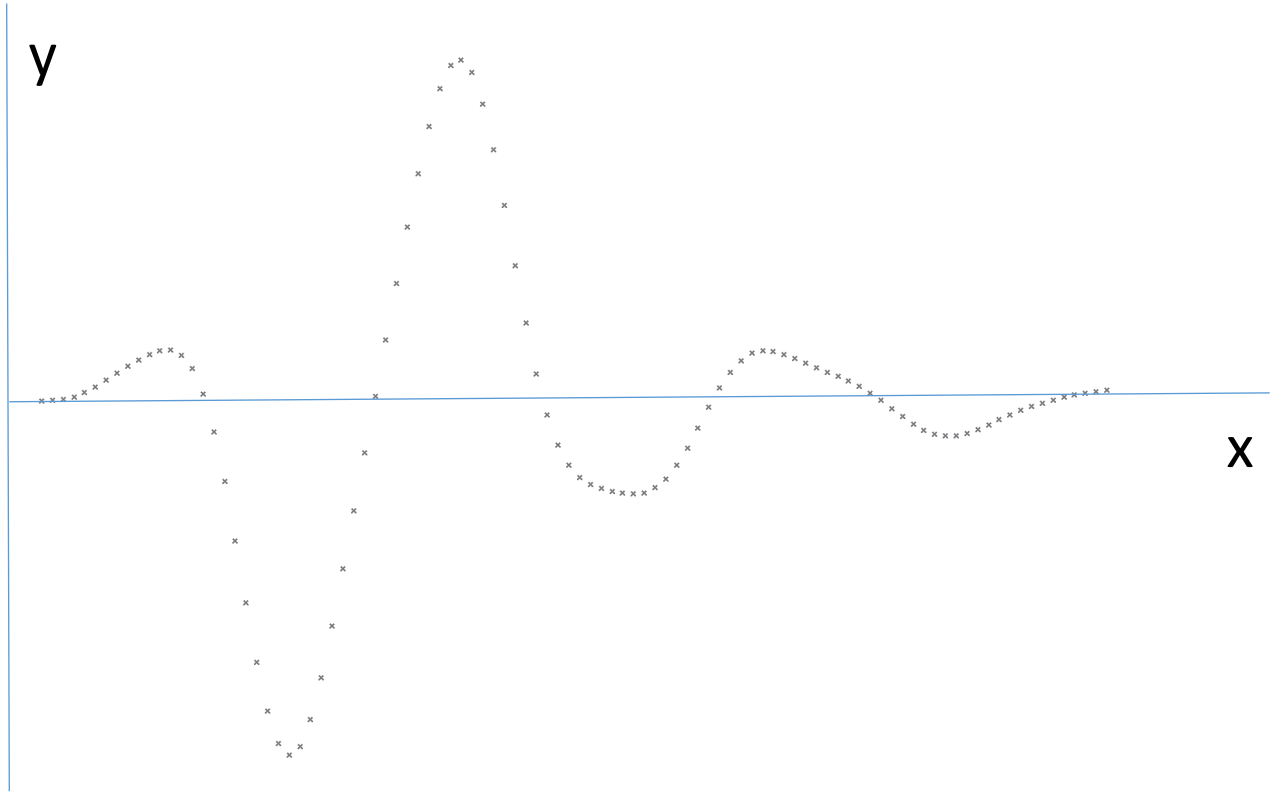
Input: Game state

Output: Score

F(X) ?

ML Examples

Modeling



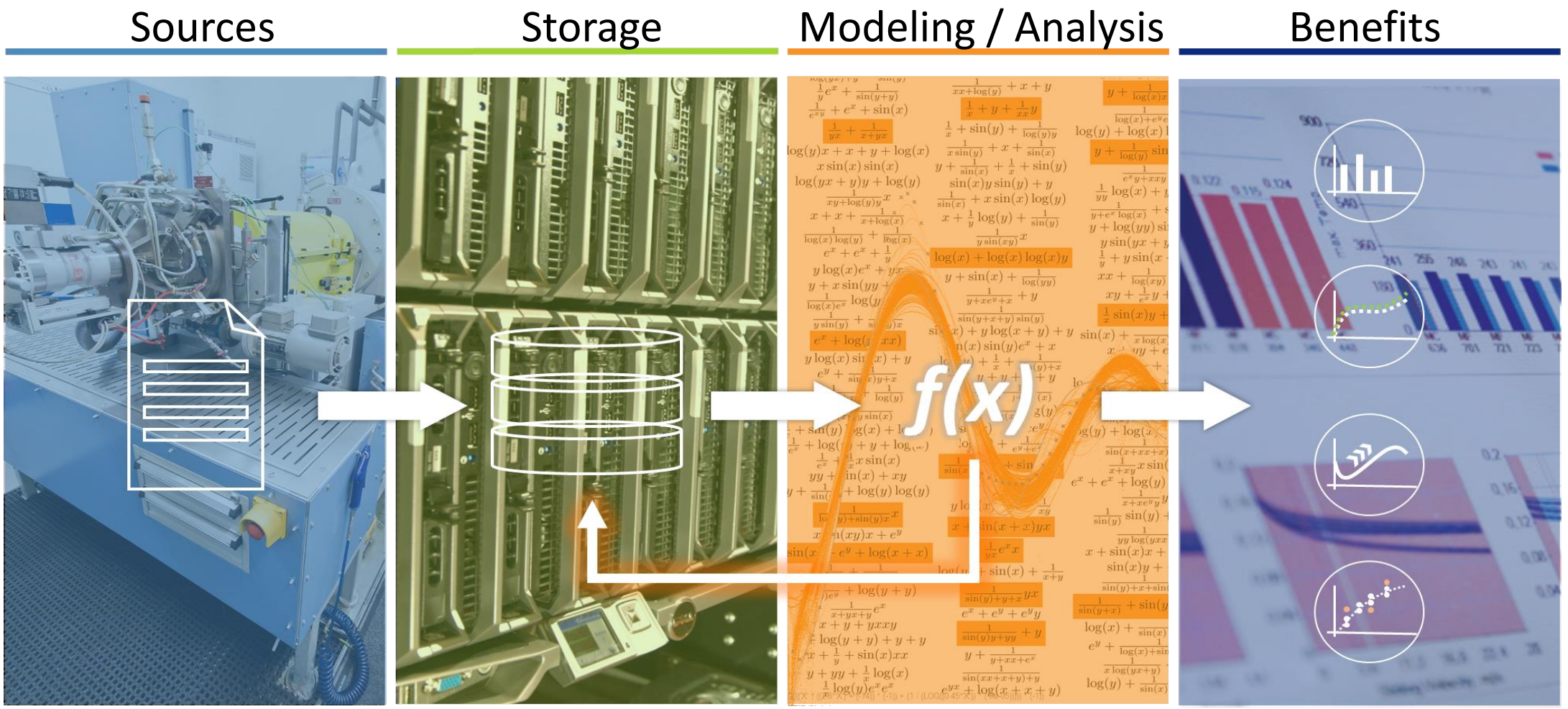
System Properties

Input: Numerical data X

Output: Predictions for Y

F(X) ?

ML Usecases



Heuristic and Evolutionary Algorithm Laboratory

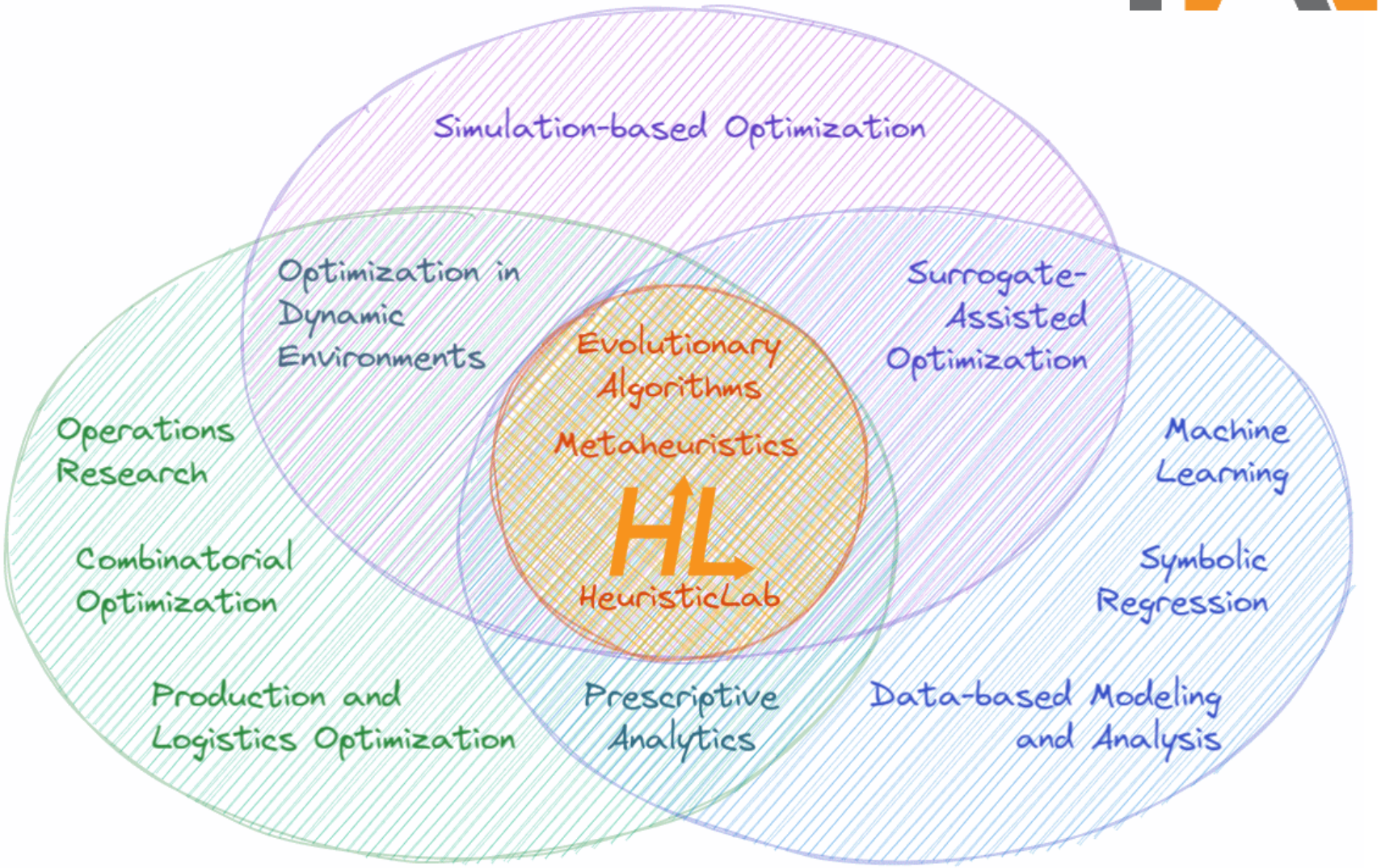
HEURISTIC AND EVOLUTIONARY ALGORITHMS LABORATORY

- Research Group
  - since 2005 at FH OÖ
  - 5 Professors
  - 20 Research associates
  - Interns, Bachelor and Master students

- Research Performance
  - > 25 research projects
  - > 200 publications
  - 11 dissertations
  - > 80 bachelor and master theses
  - various prices and awards

Research Focus

HEURISTIC AND EVOLUTIONARY ALGORITHMS LABORATORY



HeuristicLab

A Paradigm-Independent and Extensible Environment for Heuristic Optimization

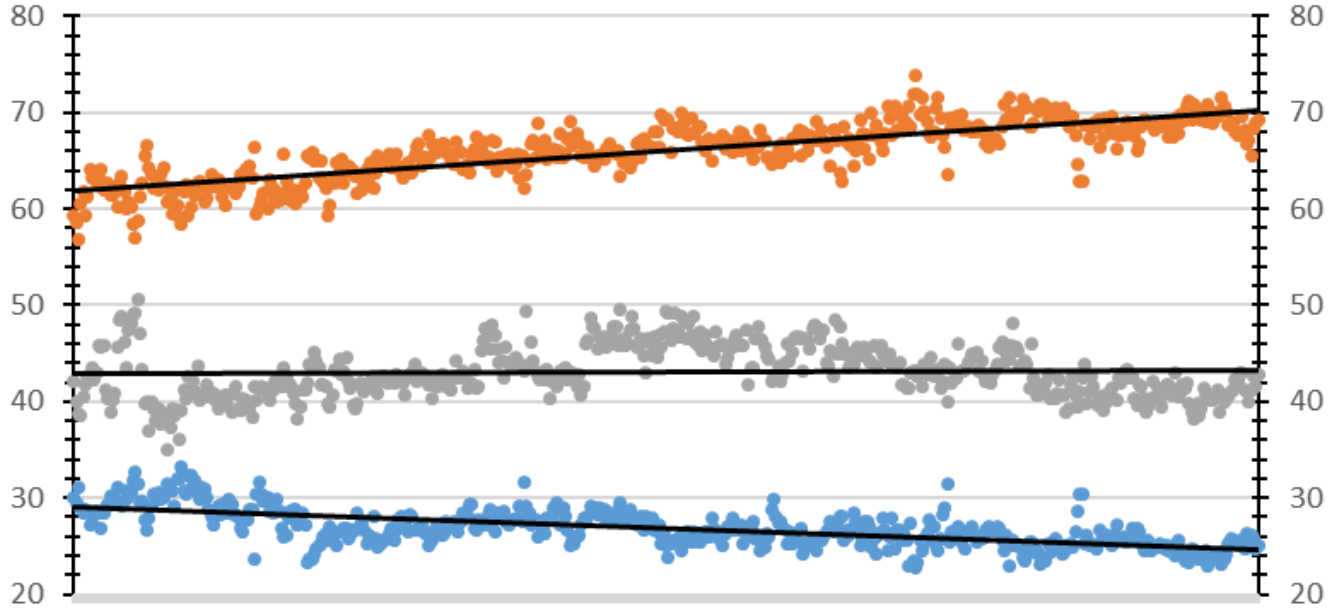
HeuristicLab is a framework for heuristic and evolutionary algorithms that is developed by members of the **Heuristic and Evolutionary Algorithms Laboratory (HEAL)** since 2002. The developers team of HeuristicLab uses this page to coordinate efforts to improve and extend HeuristicLab.

- Graphical User Interface
- Algorithm Prototyping
- Evolutionary Algorithms
- Genetic Programming
- Data Analysis
- Simulation-based Optimization
- Experiment Design and Analysis
- Plugin-based Architecture

Download HeuristicLab 3.3
Version 3.3.16, .NET4.6.1, Any CPU

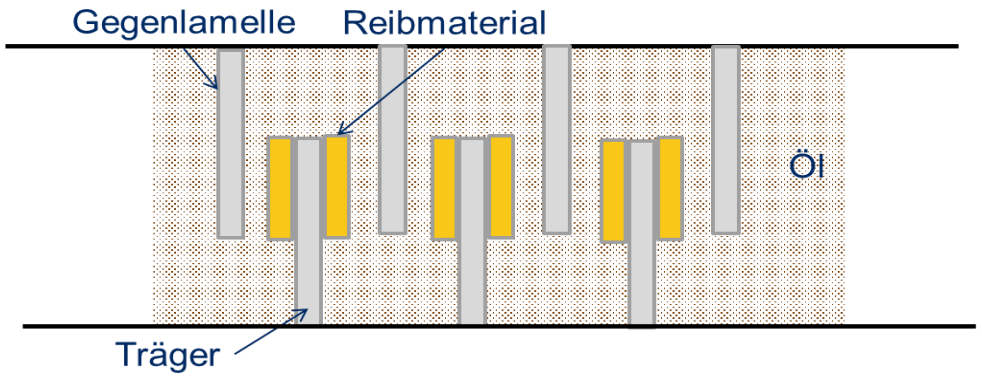
Prediction of Coke Quality

- Coke production using different coal type
- Within well-defined boundaries regarding three quality aspects

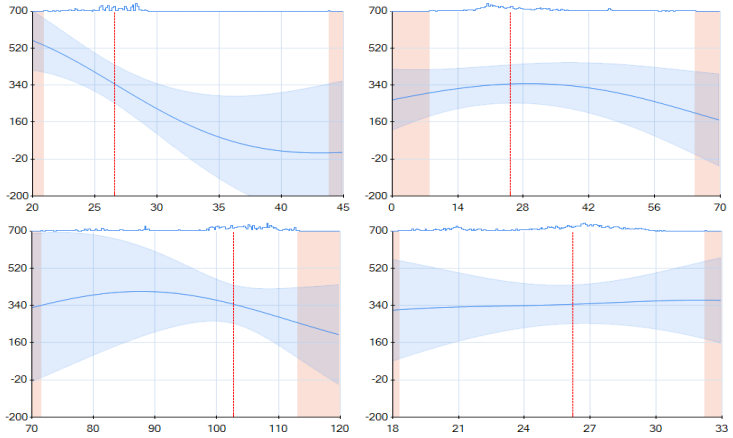


Design of Friction Systems

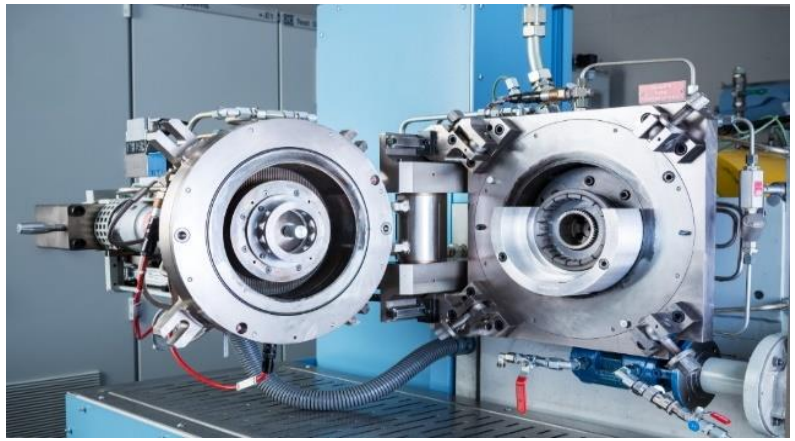
Tribological Systems



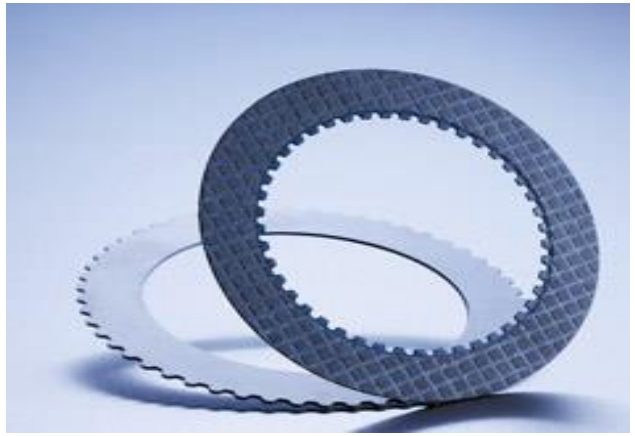
Virtual Design



Testing

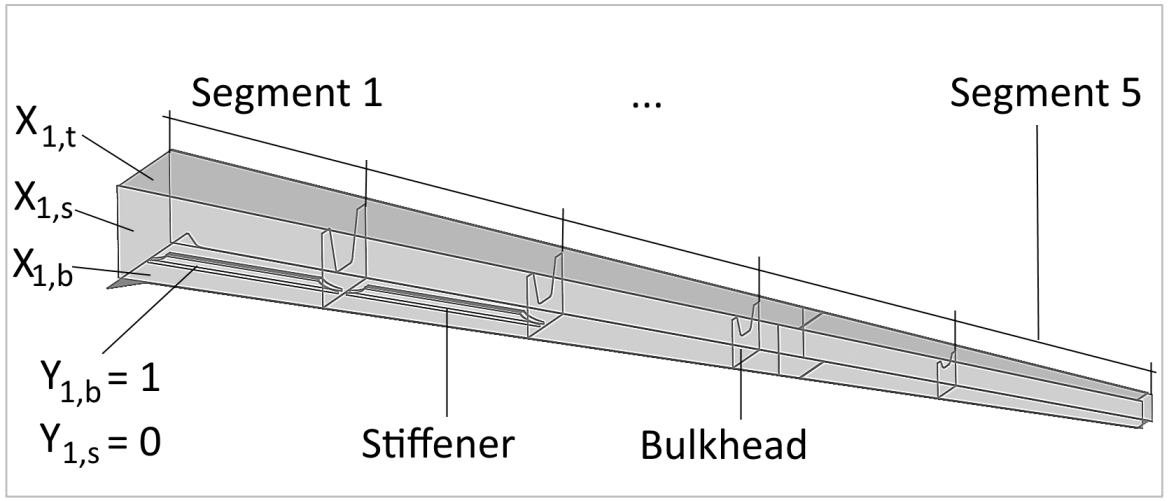


Final Product

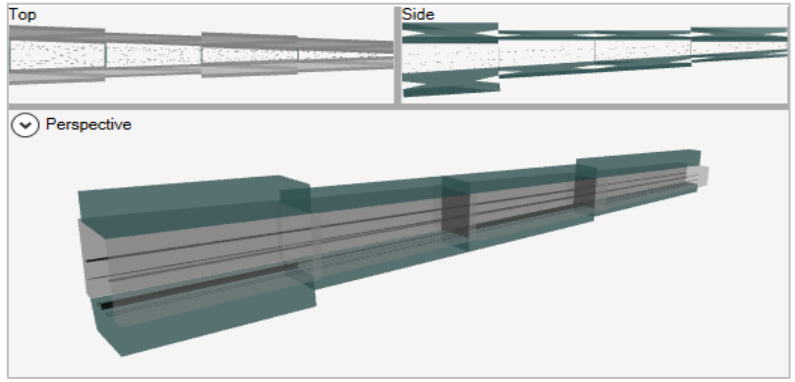


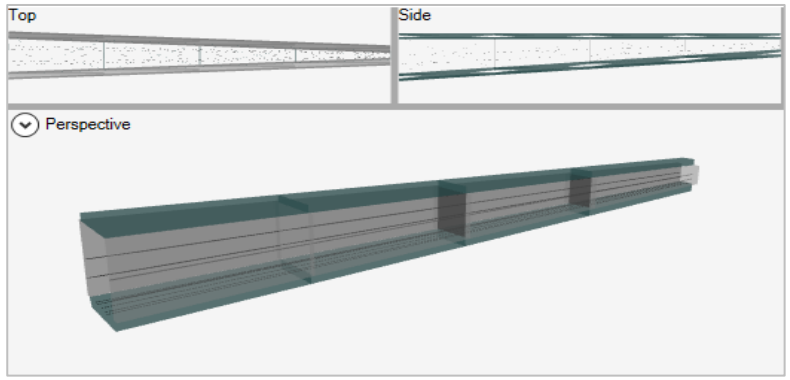
Box Type Boom Design

$$\Big(\underbrace{\overbrace{x_{1,b}, x_{1,s}, x_{1,t}}^{thicknesses}, \overbrace{y_{1,b}, y_{1,s}}^{stif\ conf}}_{Segment\ 1}, \dots, \underbrace{x_{n,b}, x_{n,s}, x_{n,t}, y_{n,b}, y_{n,s}}_{Segment\ n}, \overbrace{z_b, z_s}^{stif\ types} \Big)$$



early stage

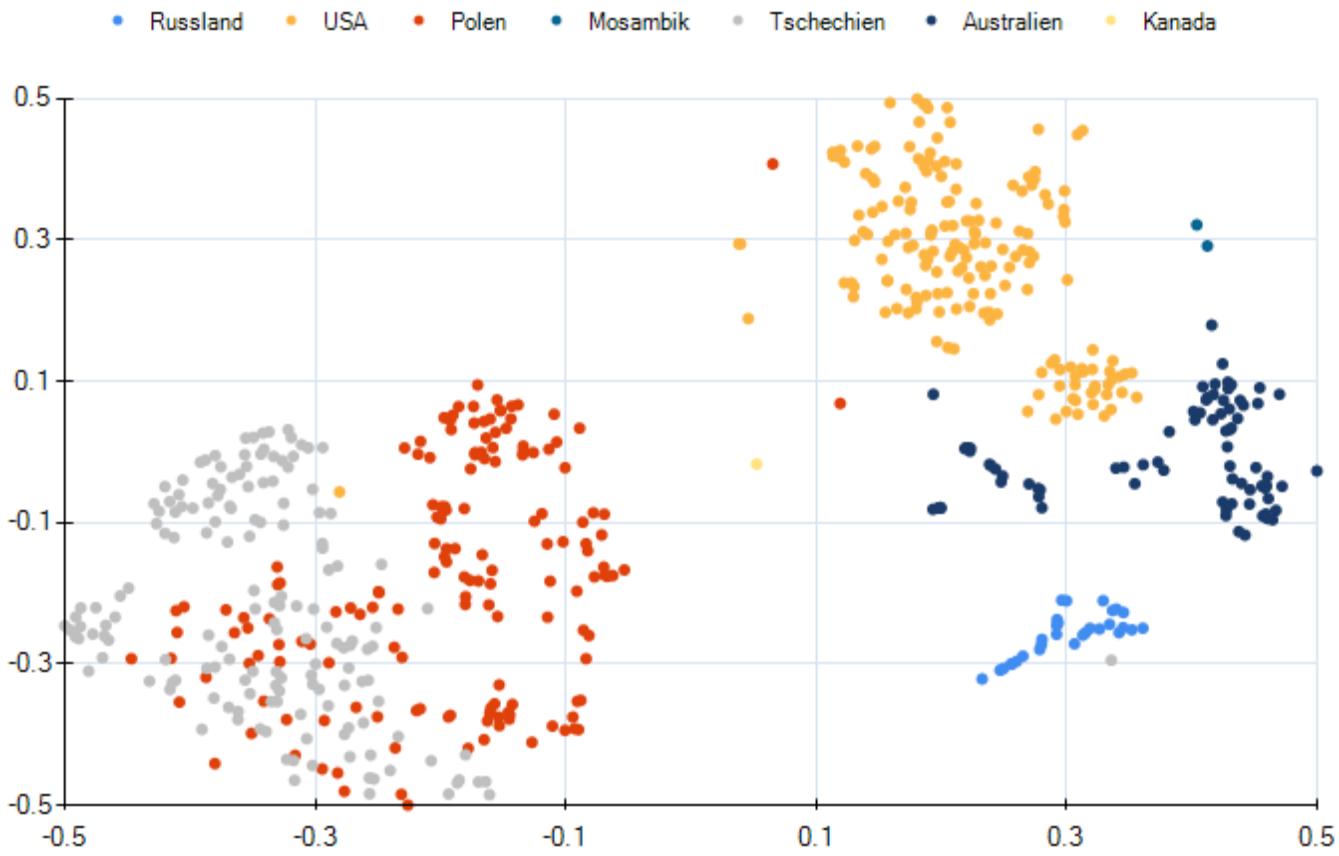




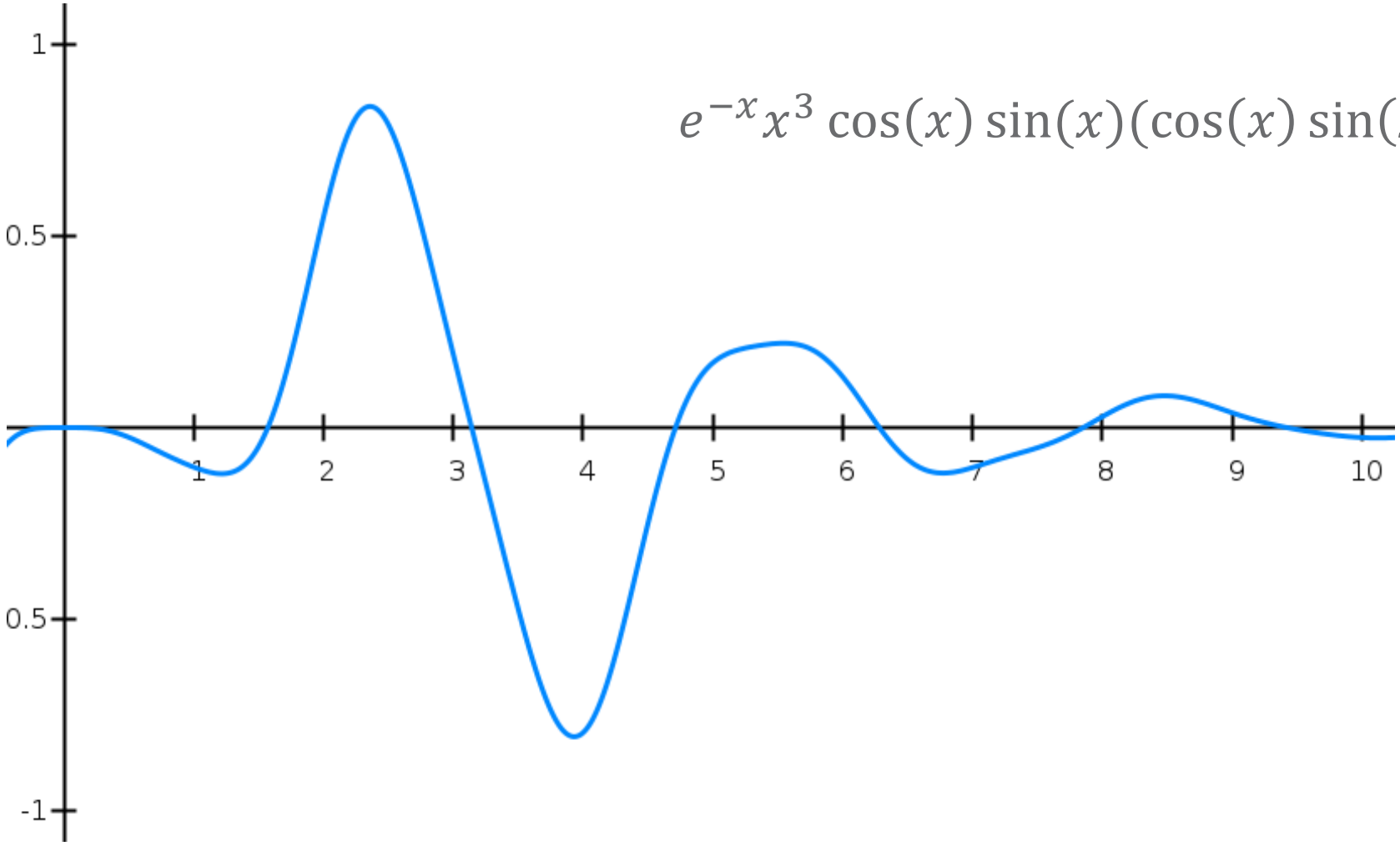
final stage

Product Origin

- Automatic chemical analysis of materials
- Control of safety relevant products
- Detection of faulty declaration

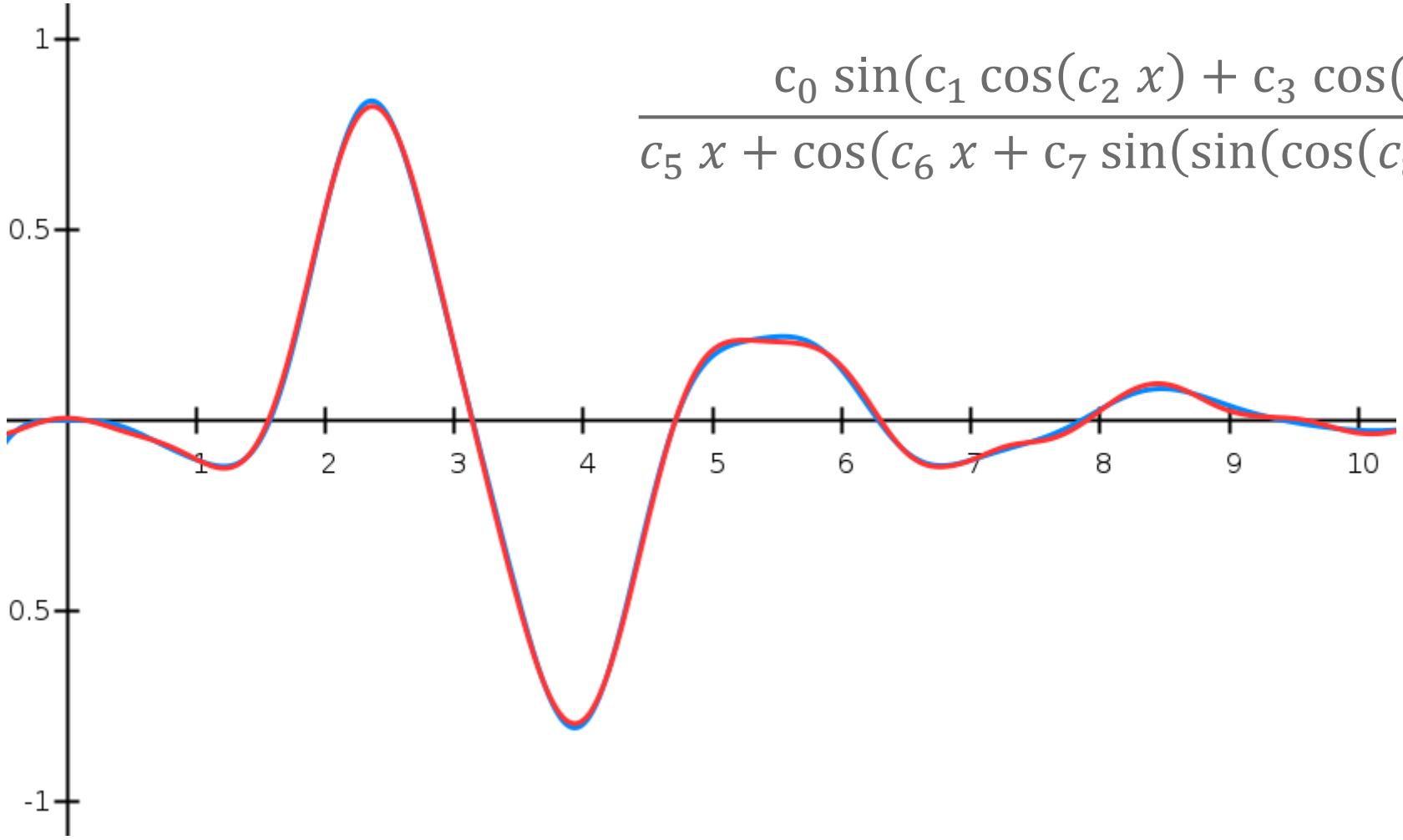


Symbolic Regression

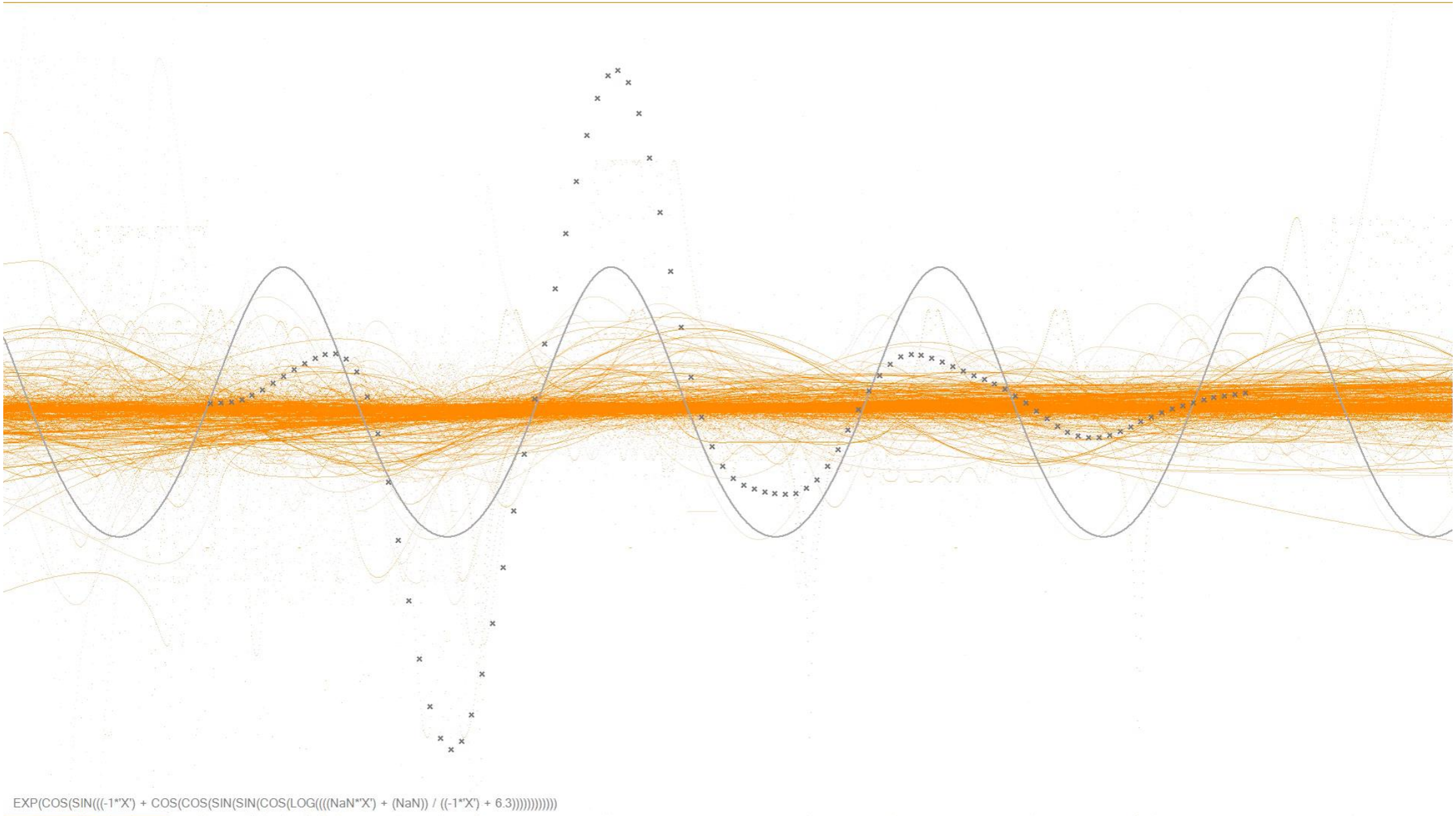


$$e^{-x}x^3\cos(x)\sin(x)(\cos(x)\sin(x)^2-1)$$

Symbolic Regression



$$\frac{c_0 \sin(c_1 \cos(c_2\, x) + c_3 \cos(c_4\, x)}{c_5\, x + \cos(c_6\, x + c_7 \sin(\sin(\cos(c_8\, x)))) + c_9}$$


EXP(COS(SIN(((-1*'X') + COS(COS(SIN(SIN(COS(LOG((((NaN*'X') + (NaN)) / ((-1*'X') + 6.3)))))))))))