Machine Learning in Physics: Introduction

Sadegh Raeisi

Outline

- Why ML?
- ML in Physics
- Types of ML

Why ML?

It’s best to review some examples.

ML in Science: Protein Folding

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › articles › article

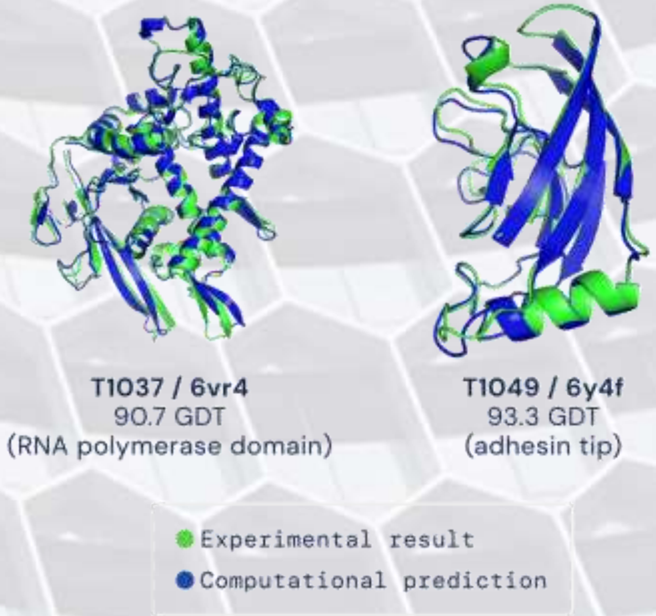
Article | Open Access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

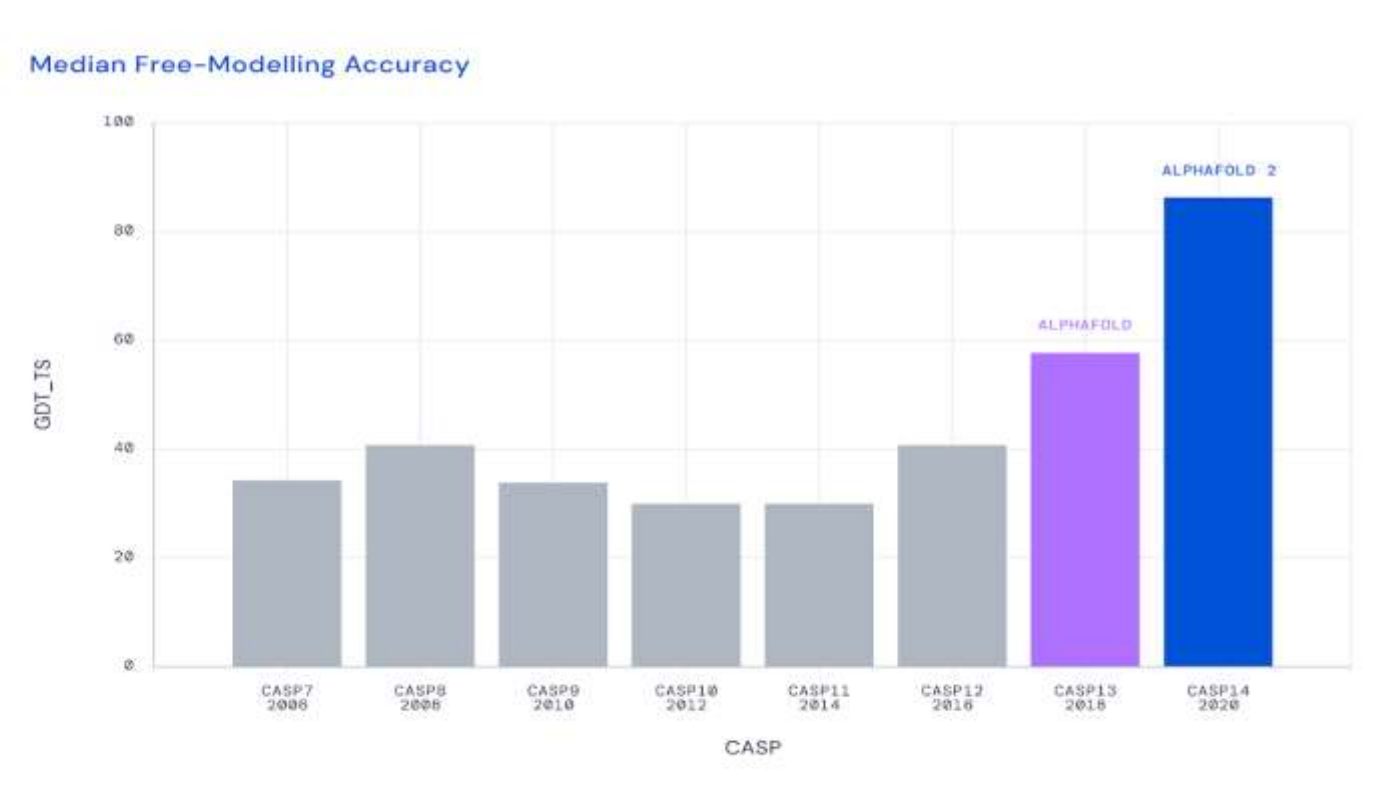
John Jumper, Richard Evans, [...]Demis Hassabis

Nature **596**, 583–589 (2021) | Cite this article

336k Accesses | **2765** Altmetric | Metrics



ML in Science: Protein Folding



ML in Science: Drug discovery

nature communication

Explore content ⌄ About the journal

nature › nature communications › articl

Article | Open Access | Published: 12 July 2

Artificial intelligence g protective therapy in i

Debashis Sahoo, Lee Swanson, Ibrahim M Rama F. Pranadinata, Courtney Tindle, Macl Sandborn, Soumita Das & Pradipta Ghos

Nature Communications **12**, Article numbe

5285 Accesses | **157** Altmetric | Metrics

nature reviews gastroenterology

Explore content ⌄ About the journal ⌄ Publish with

nature › nature reviews gastroenterology & hepatology › r

Review Article | Published: 03 January 2020

Harnessing big 'omics' data and in hepatocellular carcinoma

Bin Chen, Lana Garmire, Diego F. Calvisi, Mei-Sze Chua, Rot

Nature Reviews Gastroenterology & Hepatology **17**, 238–251 (

4354 Accesses | **24** Citations | **48** Altmetric | Metrics

nature biomedical engineering

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature › nature biomedical engineering › articles › article

Article | Published: 15 April 2021

Optimization of therapeutic antibodies by predicting antigen specificity from antibody sequence via deep learning

Derek M. Mason, Simon Friedensohn, Cédric R. Weber, Christian Jordi, Bastian Wagner, Simon M. Meng, Roy A. Ehling, Lucia Bonati, Jan Dahinden, Pablo Gainza, Bruno E. Correia & Sai T. Reddy

Nature Biomedical Engineering **5**, 600–612 (2021) | Cite this article

4788 Accesses | **2** Citations | **145** Altmetric | Metrics

ML in Science: Mathematics

NewScientist

Volume 242, Issue 3228, 4 May 2019, Page 9

News & Technology

Machine learning

Google's AI mathematician

Leah Crane

HOList: An Environment for Machine Learning of Higher-Order Theorem Proving

Kshitij Bansal, Sarah M. Loos, Markus N. Rabe, Christian Szegedy, Stewart Wilcox

We present an environment, benchmark, and deep learning driven automated theorem prover for higher-order logic. Higher-order interactive theorem provers enable the formalization of arbitrary mathematical theories and thereby present an interesting, open-ended challenge for deep learning. We provide an open-source framework based on the HOL Light theorem prover that can be used as a reinforcement learning environment. HOL Light comes with a broad coverage of basic mathematical theorems on calculus and the formal proof of the Kepler conjecture, from which we derive a challenging benchmark for automated reasoning. We also present a deep reinforcement learning driven automated theorem prover, DeepHOL, with strong initial results on this benchmark.

Comments: Accepted at ICML 2019

Subjects: **Logic in Computer Science (cs.LO)**; Artificial Intelligence (cs.AI); Machine Learning (cs.LG)

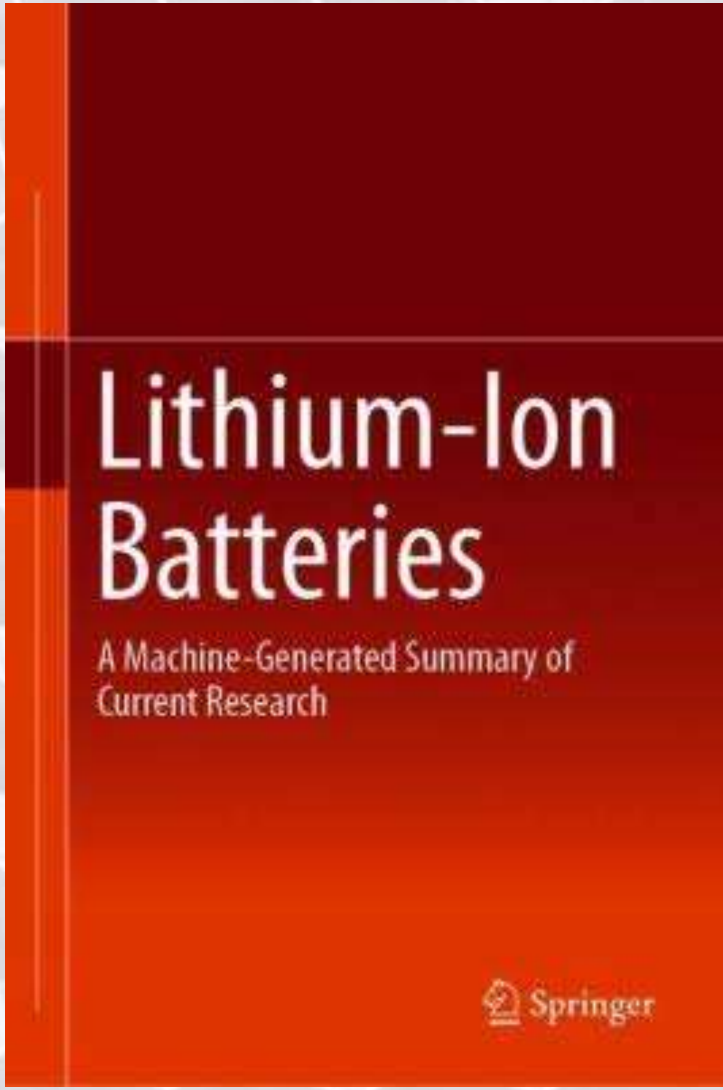
Cite as: arXiv:1904.03241 **[cs.LO]**

(or arXiv:1904.03241v3 **[cs.LO]** for this version)

Artificial intelligence trained by google learns to prove 1200 theorems.

[1904.03241] HOList: An Environment for Machine Learning of Higher-Order Theorem Proving (arxiv.org)

ML in Science: writing books



Can machines replace scientists?

What are the key aspects that cannot be replaced?

ML in Physics

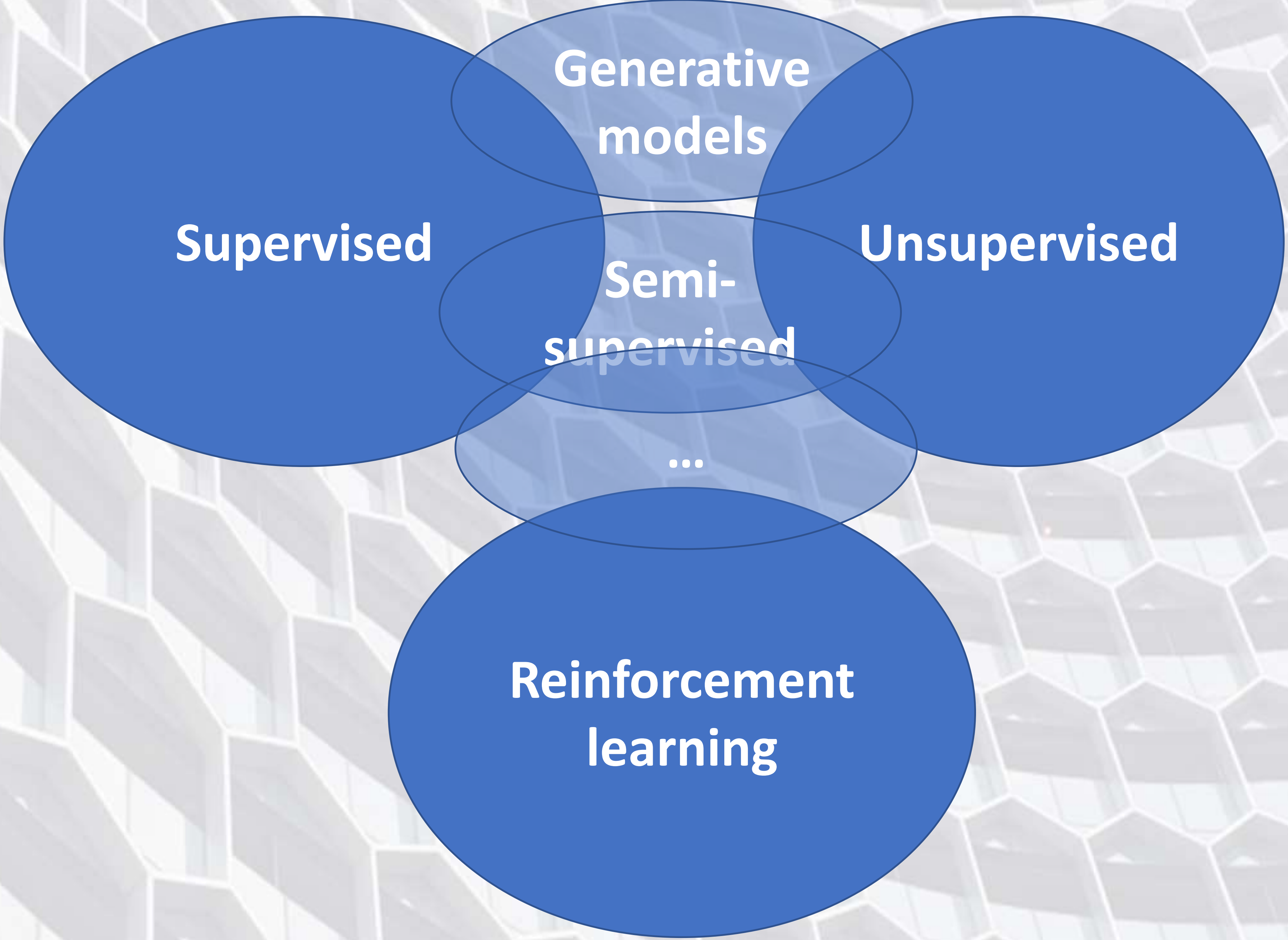
ML in Physics

Physics

- Statistical concepts
- Models (e.g. energy based models, Boltzmann Machines, …)
- New training algorithms

ML

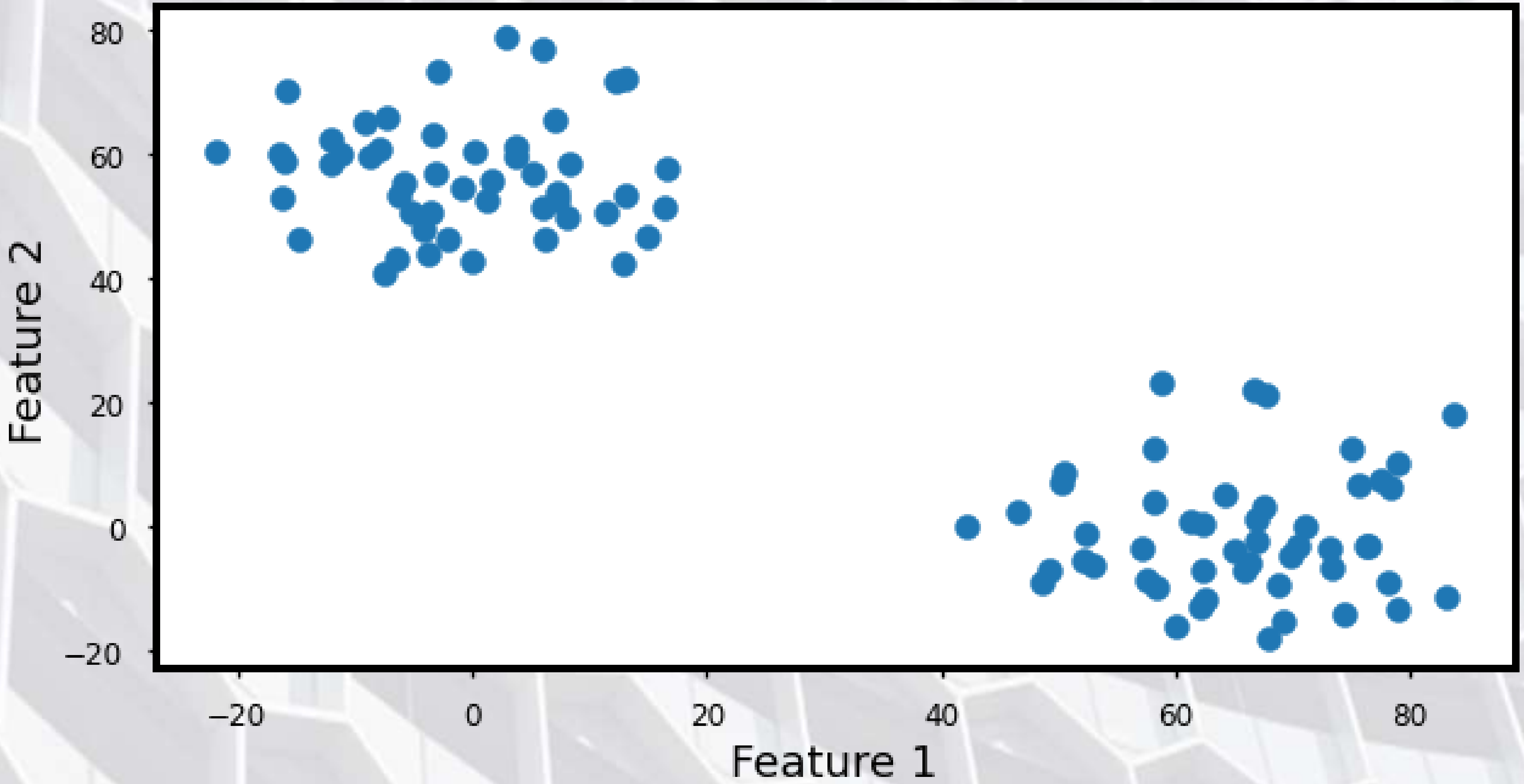
Types of ML


Generative models
Supervised
Semi-supervised
Unsupervised
...
Reinforcement learning

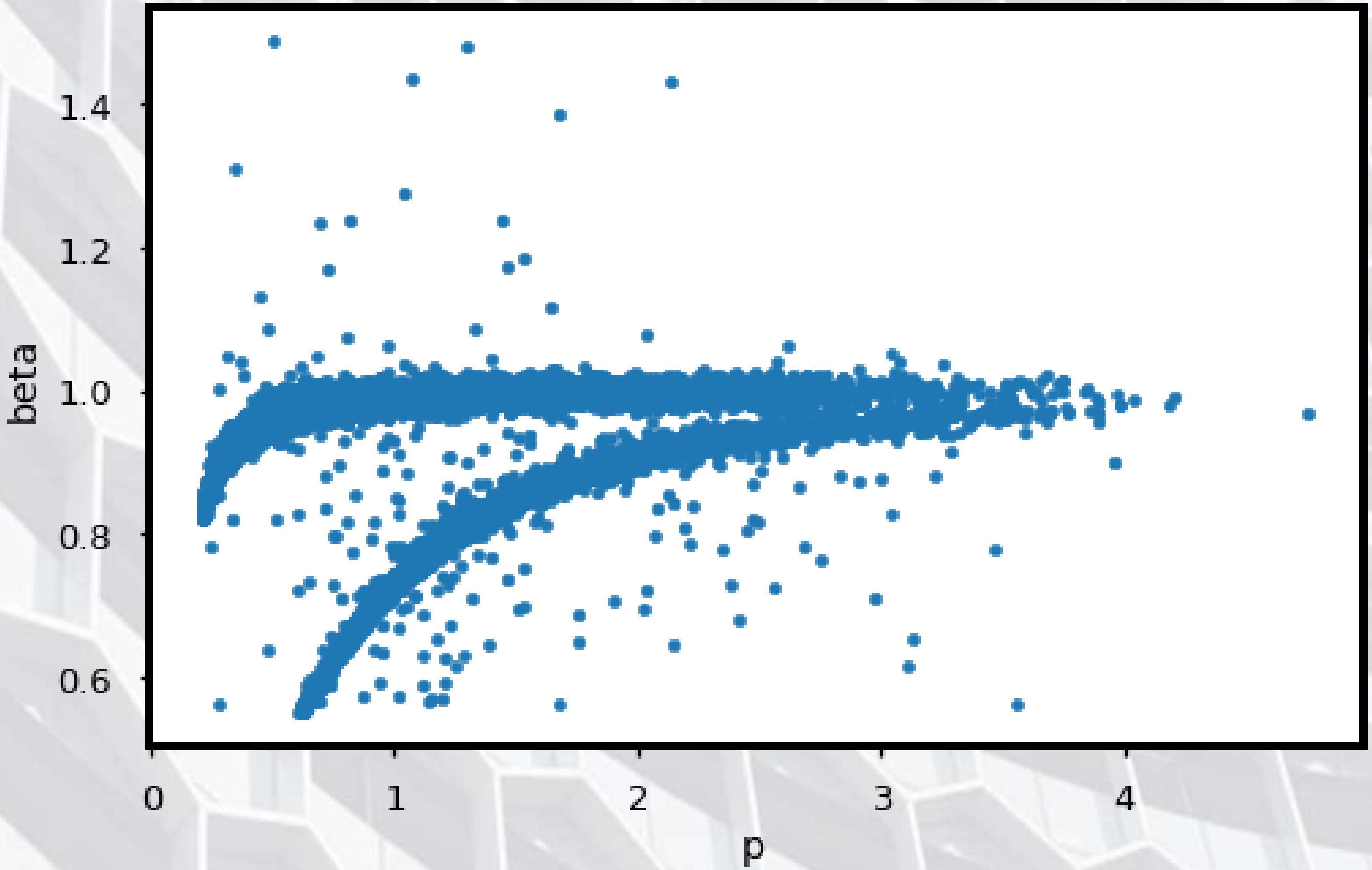
Unsupervised

- Clustering
- Dimensionality reduction
- Anomaly detection
- …

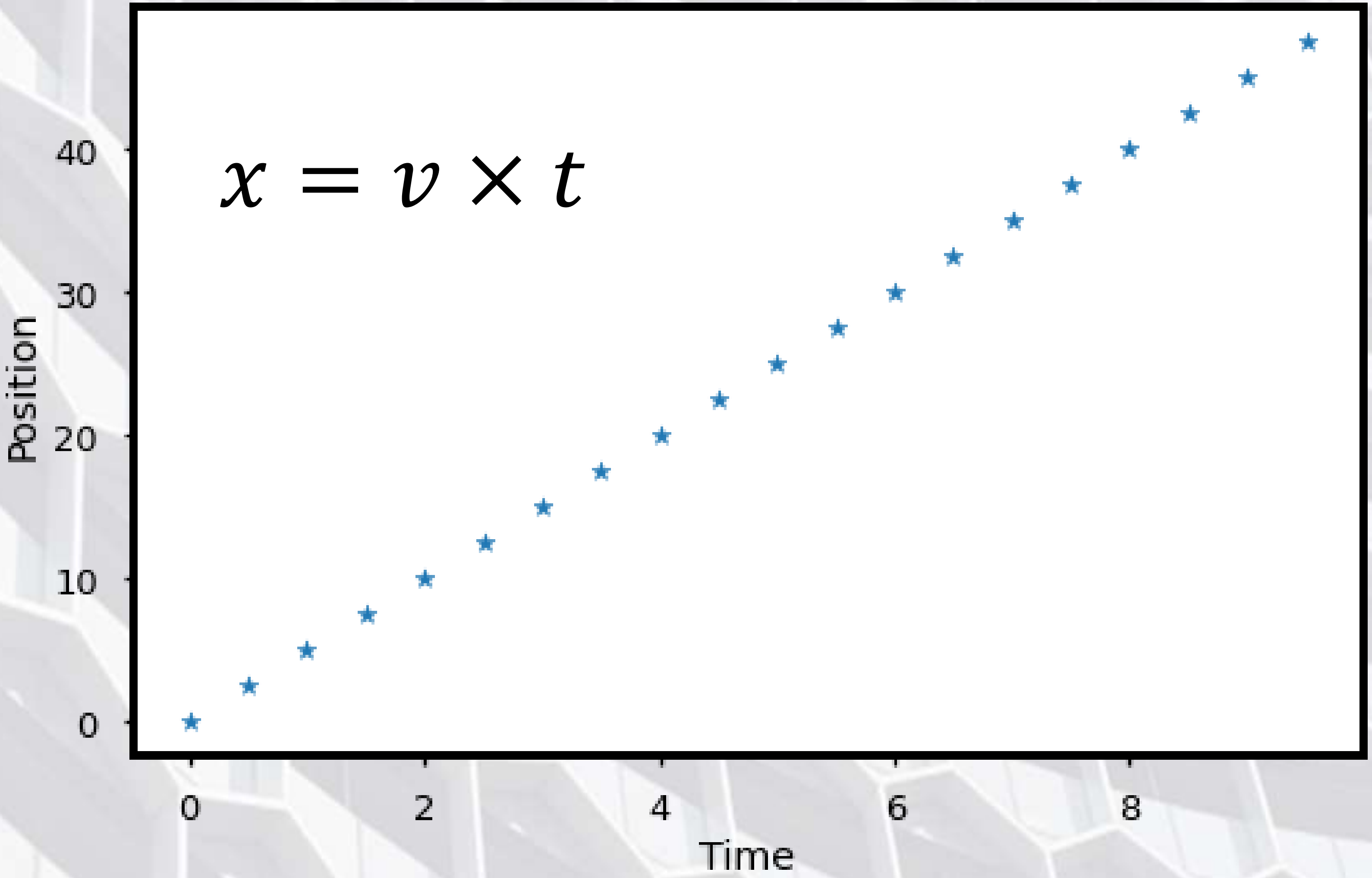
Unsupervised: Clustering



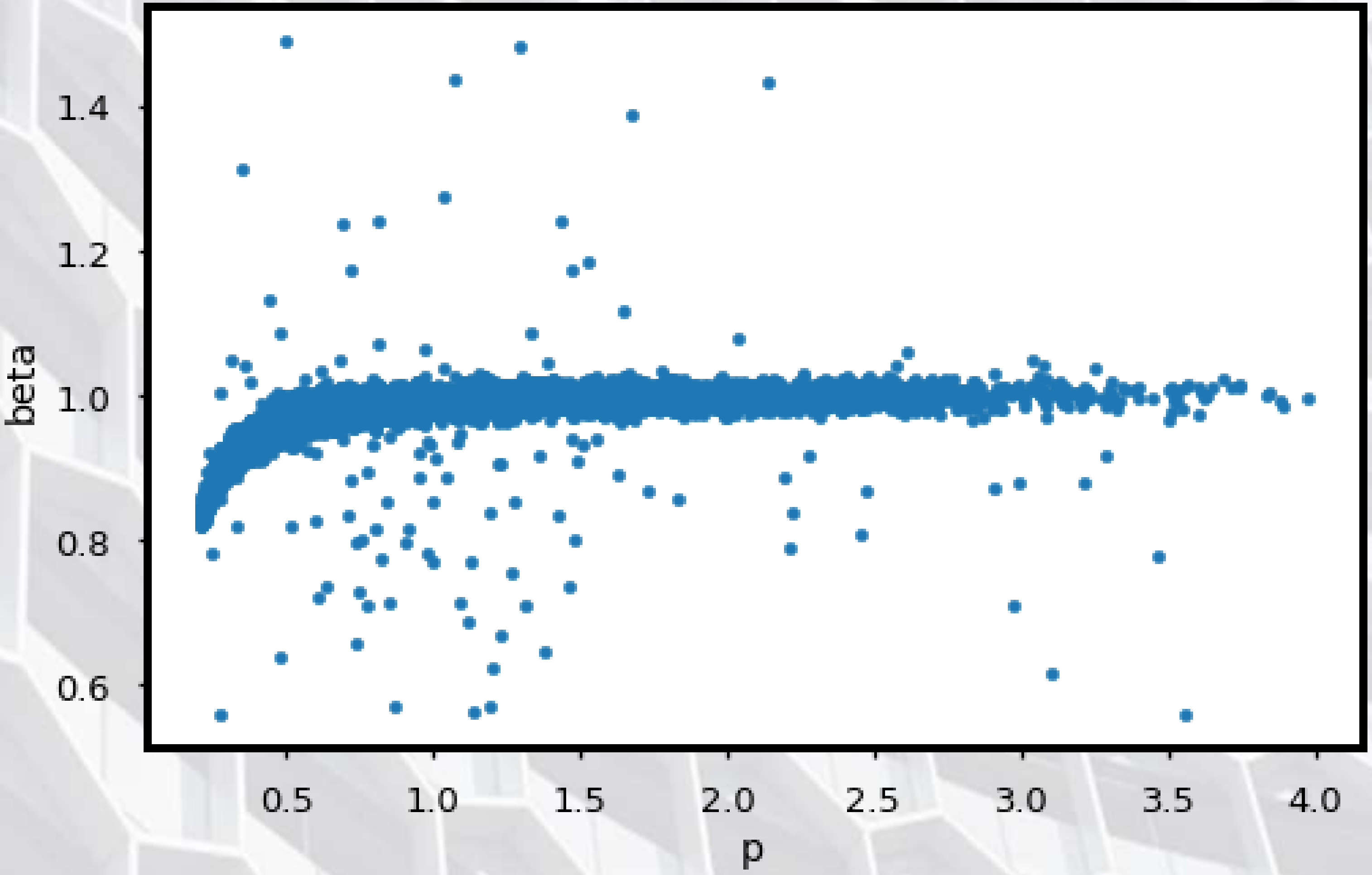
Unsupervised: Clustering



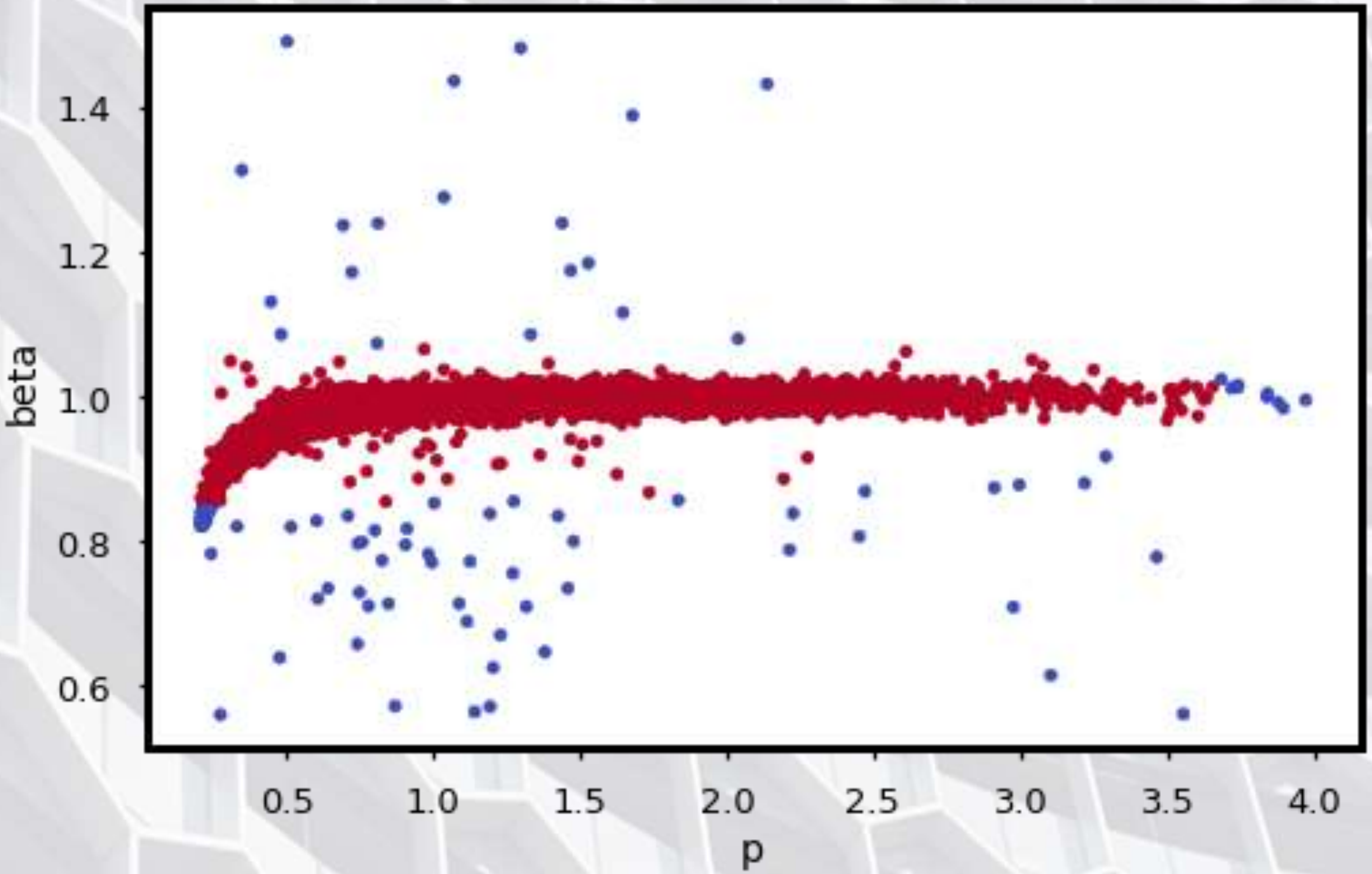
Unsupervised: Dimensionality reduction



Unsupervised: Anomaly detection



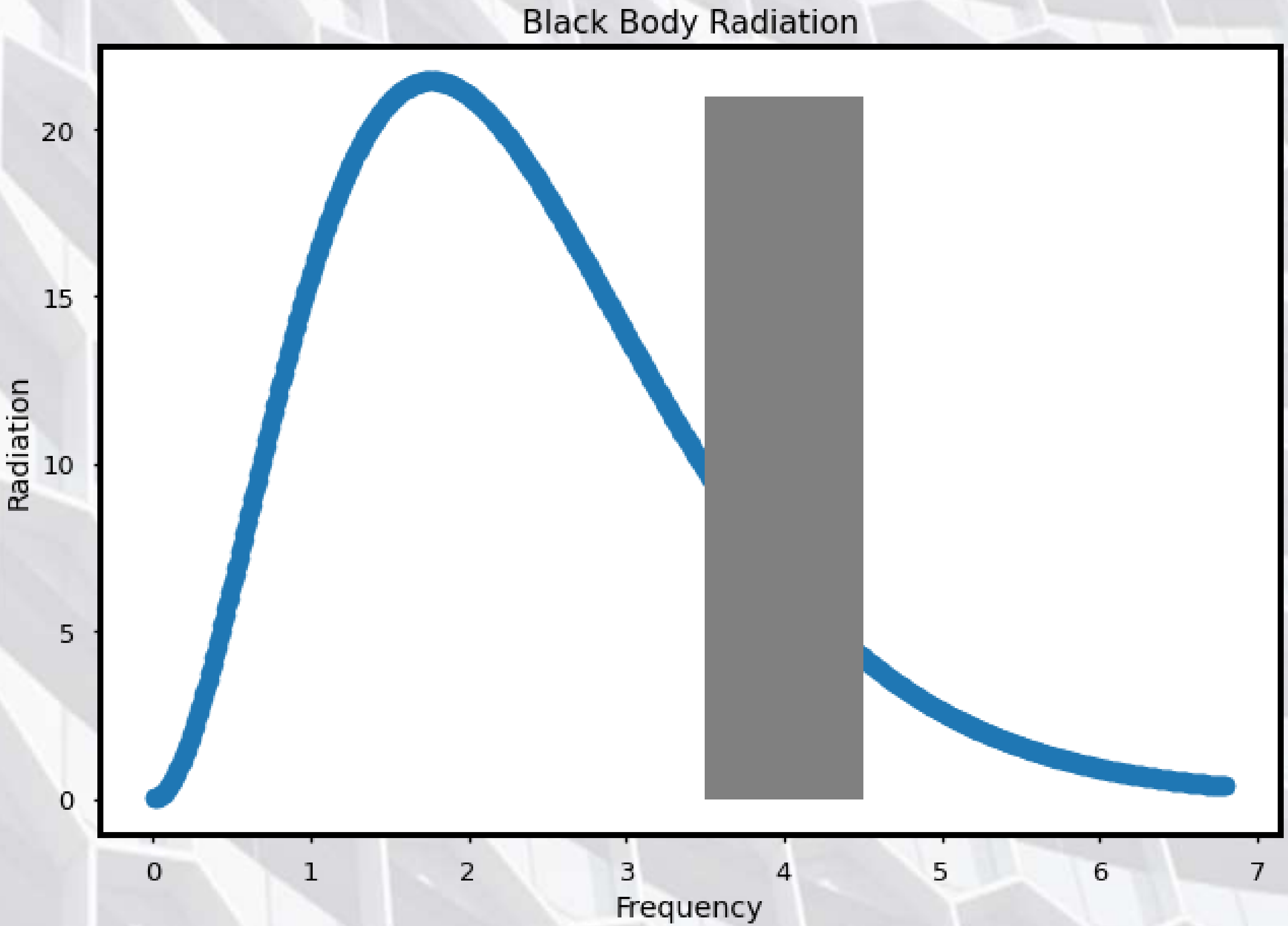
Unsupervised: Anomaly detection



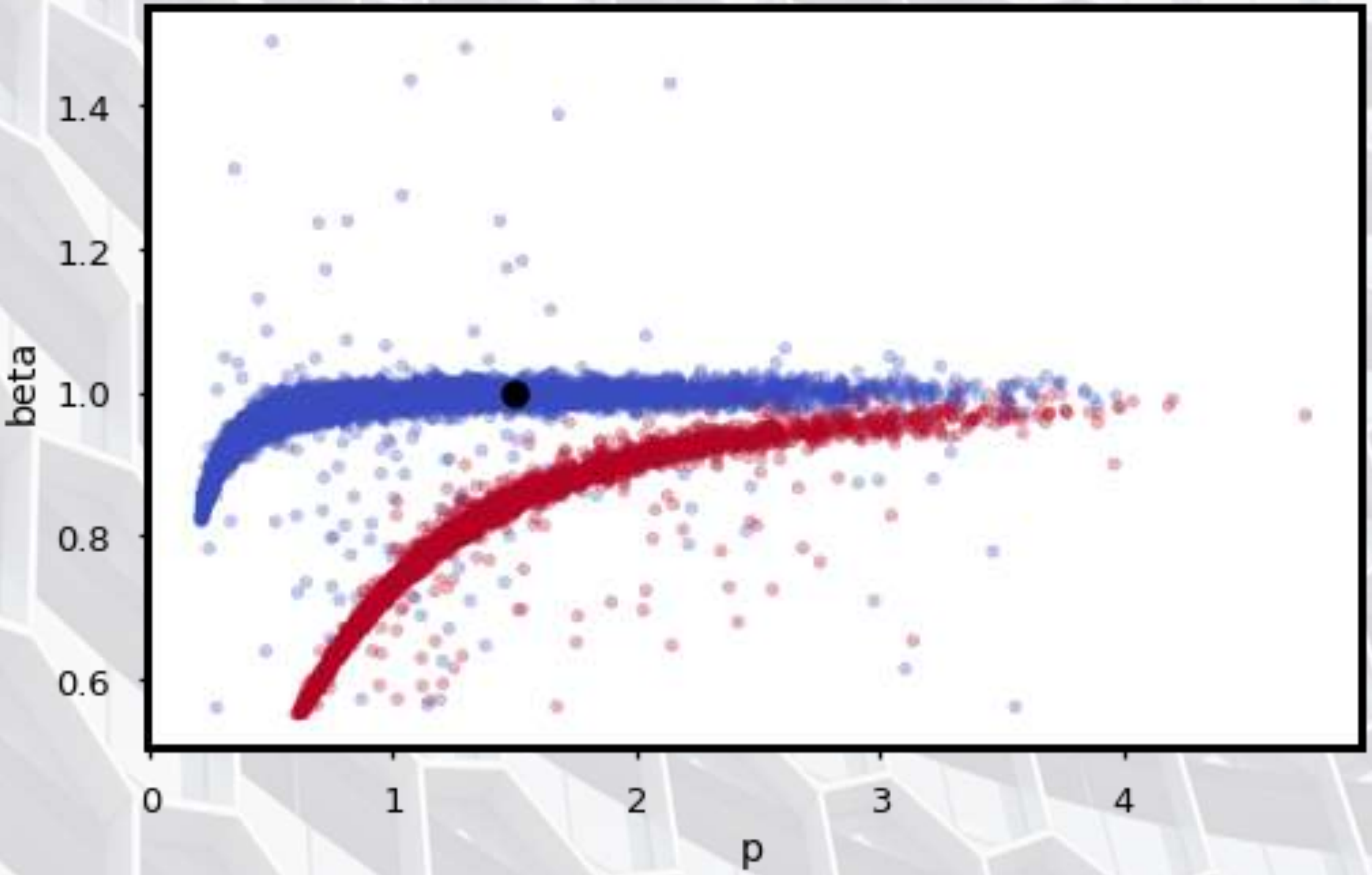
Supervised

- Regression
- Classification
- (Logistic Regression)

Supervised: Regression



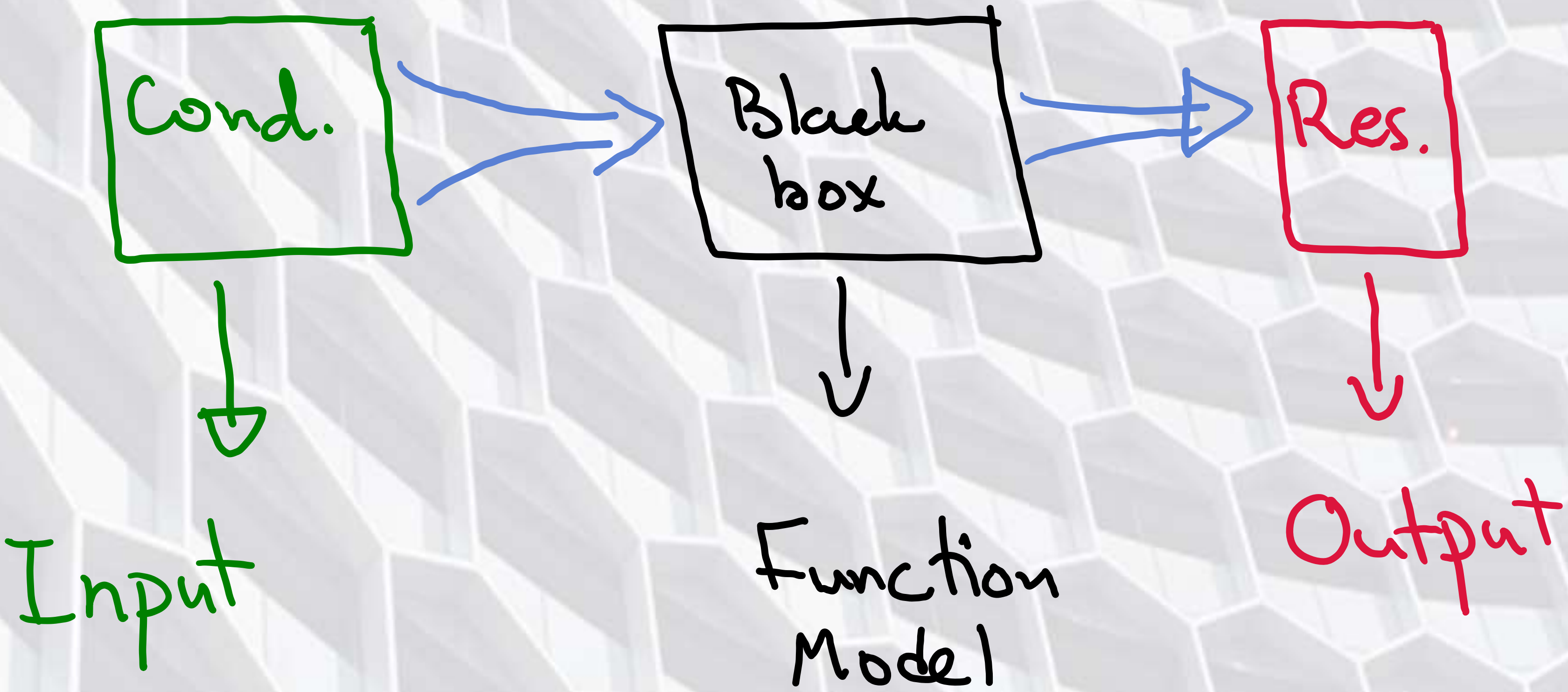
Supervised: Classification



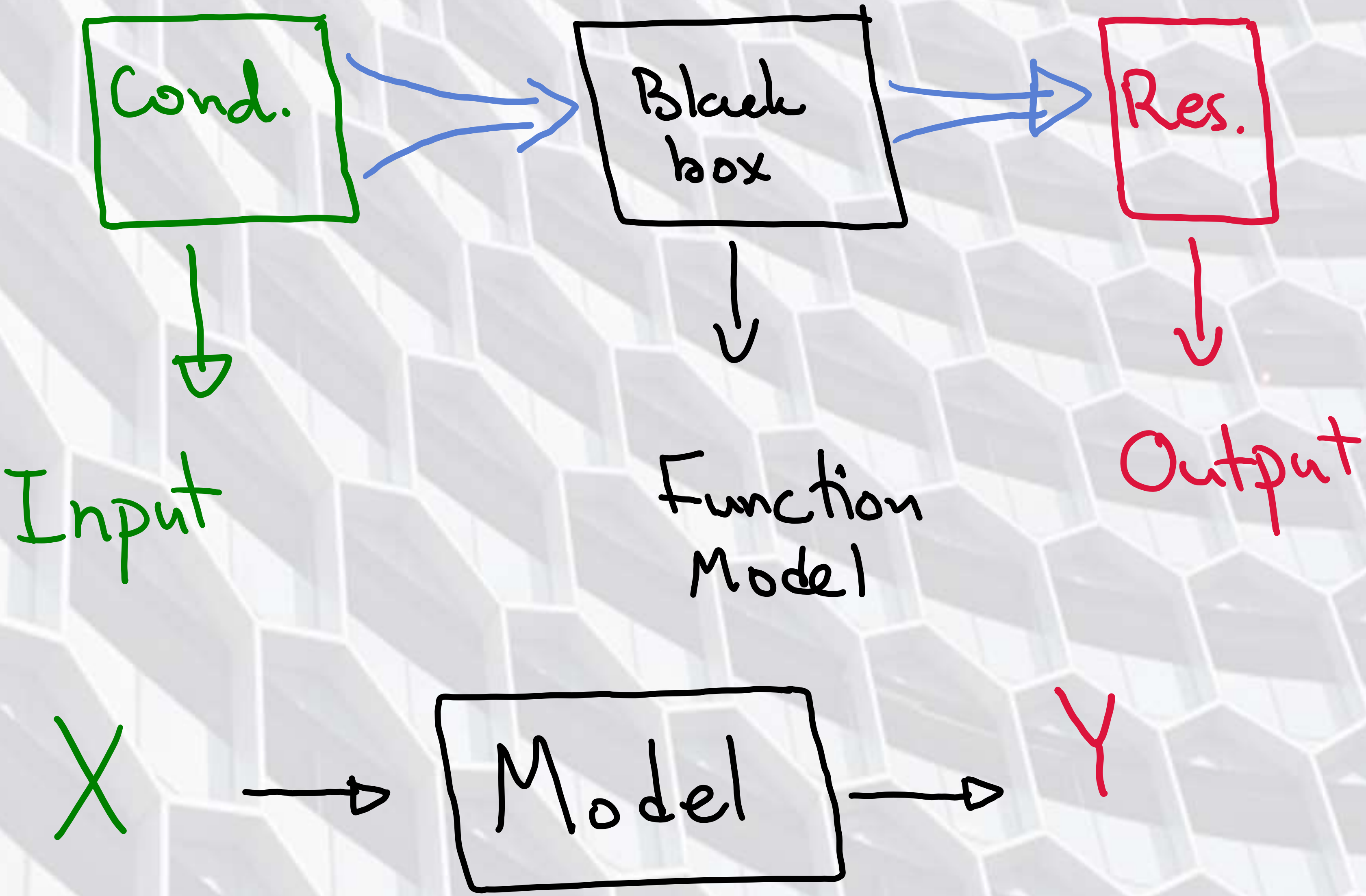
When do we use ML (supervised)

- We don’t know how something works:
  - Complex systems with too many variables: weather …
  - Complex systems with complex functionality
  - Simple problems that needs to be automated (boring ones)

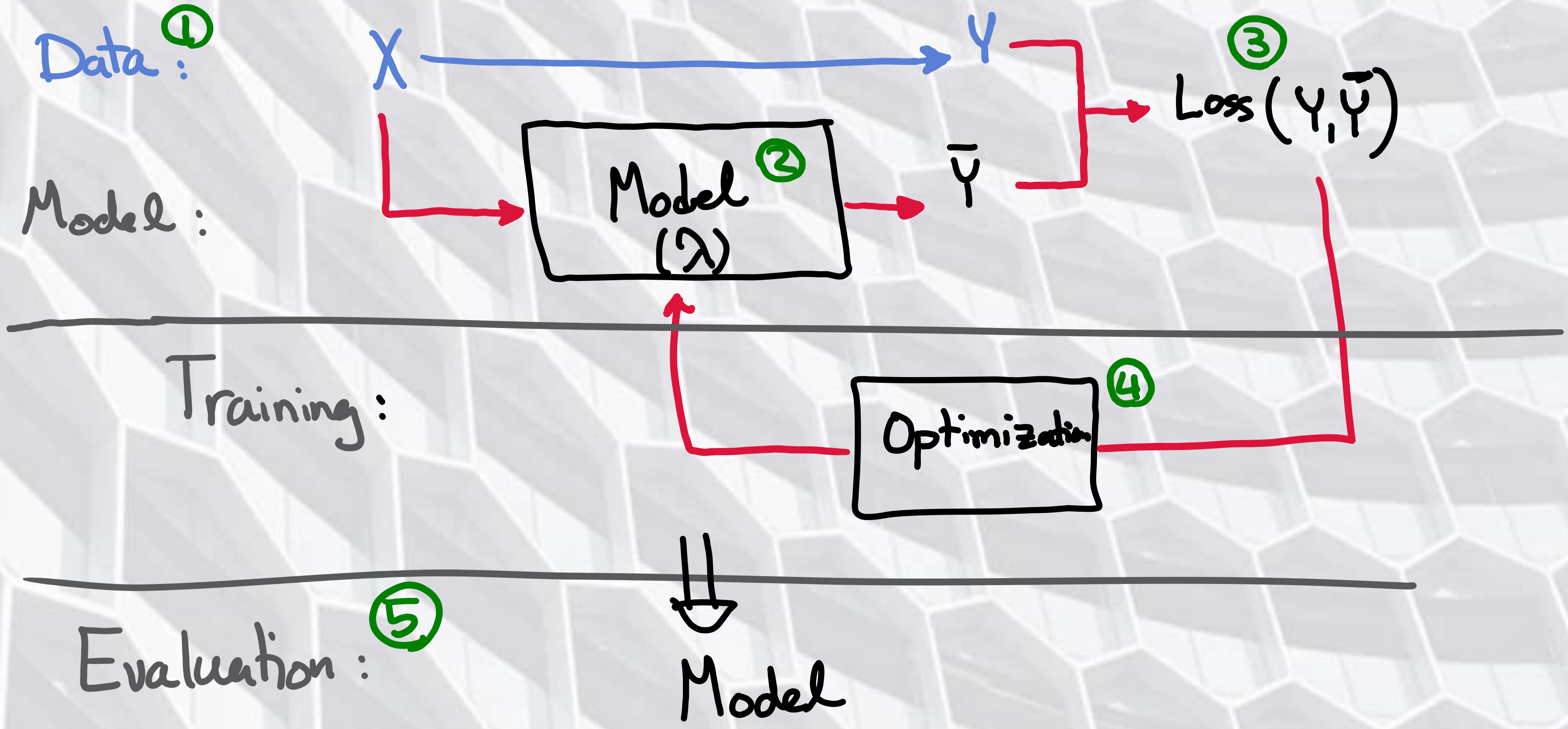
Supervised



Supervised



Supervised: Ingredients



Supervised: Ingredients

- Hypothesis (Model)
- Data
- Distance function (loss)
- Training algorithm
- Evaluation metrics

Make a model

- Make a clustering algorithm
- Make a regression algorithm
- Make a classification algorithm