



Machine Learning with Python

PXL i-Talent



Overview

 Introduction

 What is Machine Learning

 Practical exercises

About us



Dr. Glenn Cich

- Master Computer Science
- PhD Mobility Sciences
- Data/Software Engineer



Lorenz Feyen

- Master Nuclear Physics
- Data Scientist

A decorative network graph on the left side of the slide, consisting of blue dots (nodes) connected by thin blue lines (edges), forming a complex web-like structure.

What is InfoFarm?

- Data Science company
- Help customers find insights and business value in their data
- Analysis & implementations
- 7 data scientists, 6 data engineers
- Located in Hasselt, Merelbeke & Kontich



Infofarm

We harvest business value



OUR PASSION

Plow data – Grow information – Harvest value

Data Science



Applied Artificial Intelligence services,
Machine Learning & Deep Learning
techniques

Big Data



Setting up and maintaining on-premise
and Cloud-based Big Data architectures
with Apache Spark and Hadoop

Training & Workshops



Training and guiding organisations
through the digital era using the power
of Data Science

YOU
ARE
HERE

CRONOS GROEP

Powered by



What is InfoFarm?

- Part of Xplore Group

- Mobile development
- Web development
- E-commerce
- Lots of partner companies for InfoFarm

- Part of Cronos Group

- Largest independent IT services supplier in Belgium
- 6000+ Consultants



InfoFarm customers



Vlaamse
overheid



proximus

VDAB



MEDIAHUIS



LINEAS
YOUR FREIGHT FORCE





What is Machine Learning?

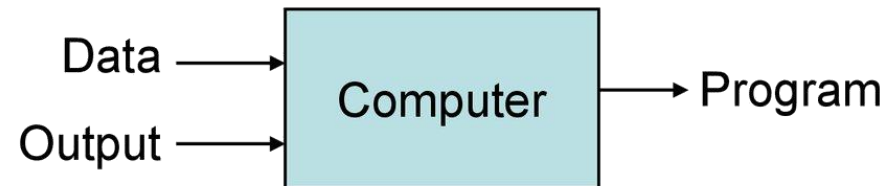
Machine Learning: Concept

- Learn and make predictions on data
- No instructions for single cases
- behavior is determined from past data





Traditional Programming



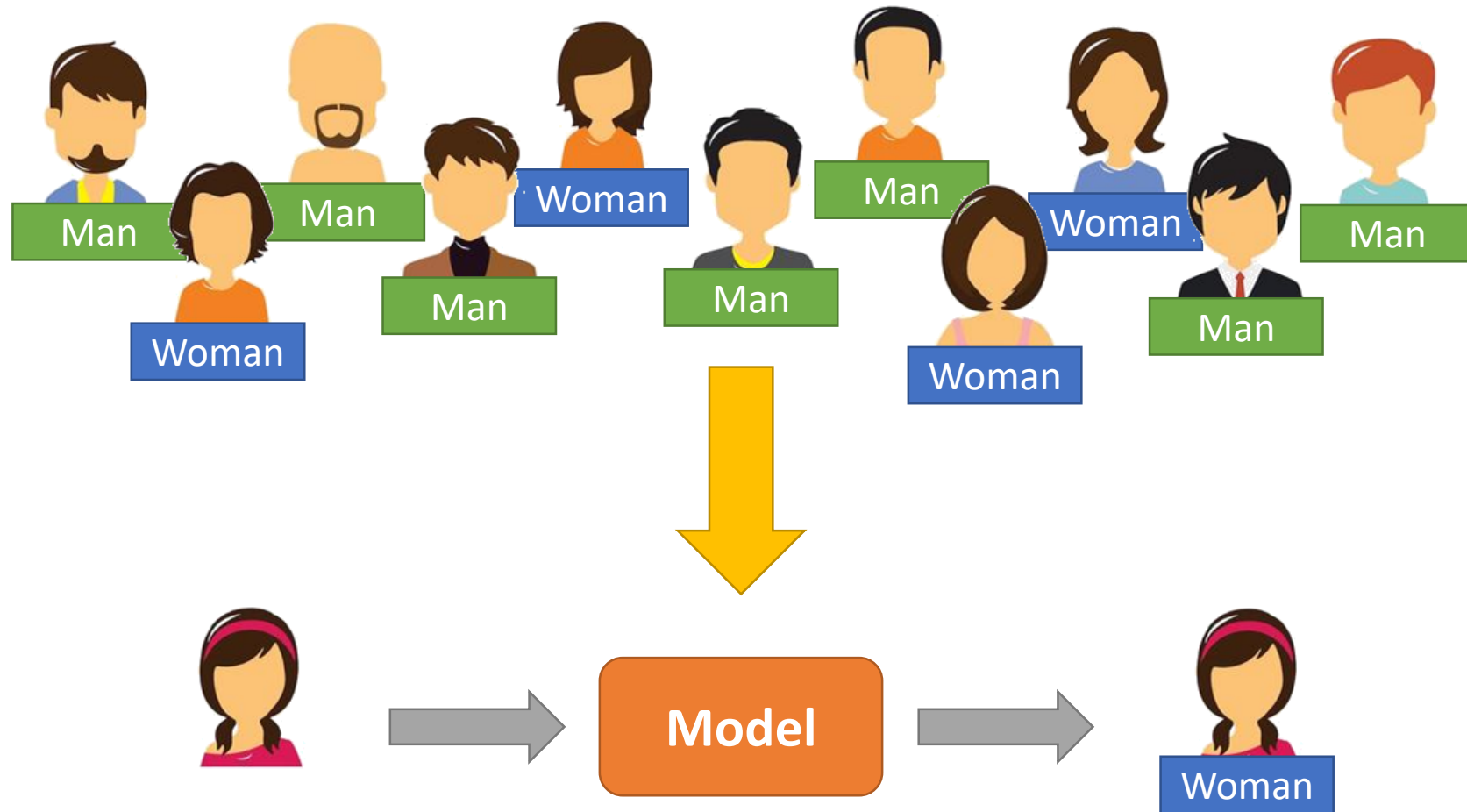
Machine Learning



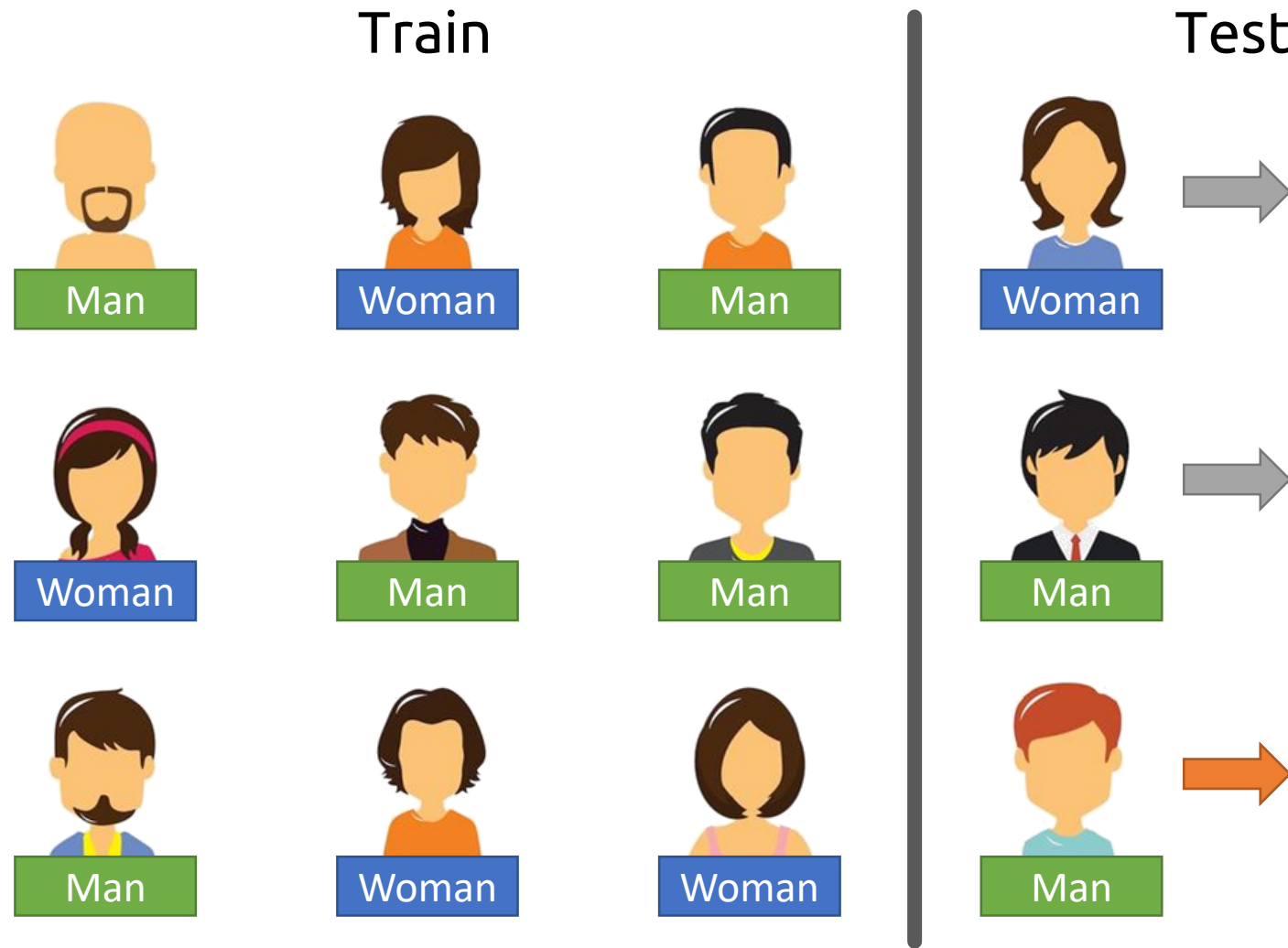
Machine Learning: Features

	Gender	Age	Height (cm)	Hair color
	Man	24	184	Black
	Woman	27	169	Brown
	Man	19	192	Red
	Woman	15	155	Brown

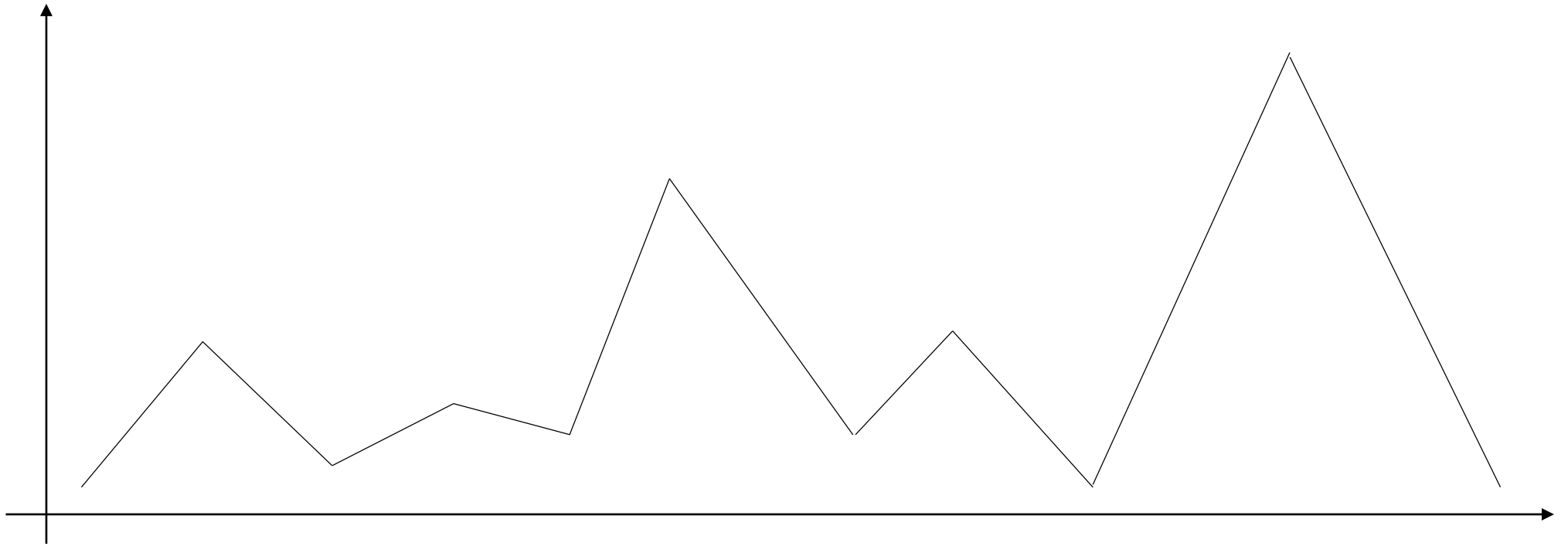
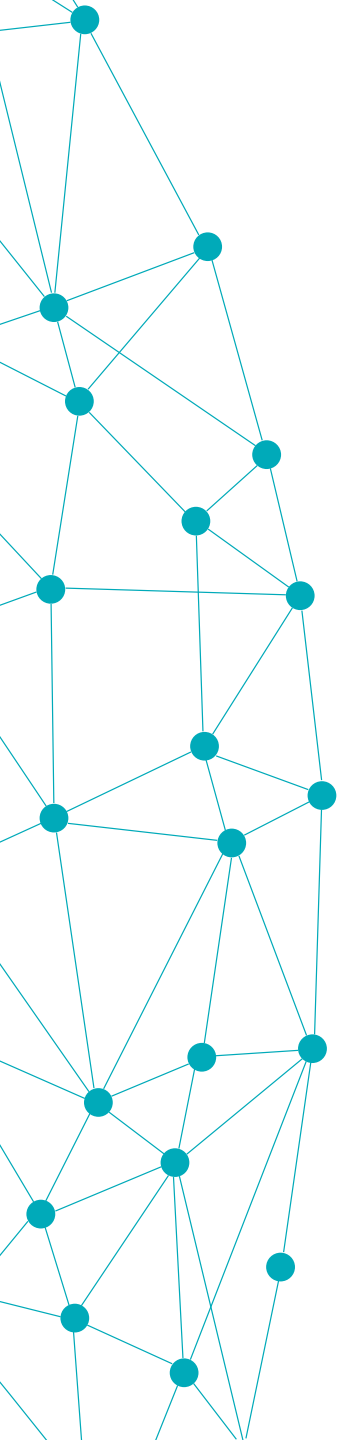
Machine Learning: Supervised



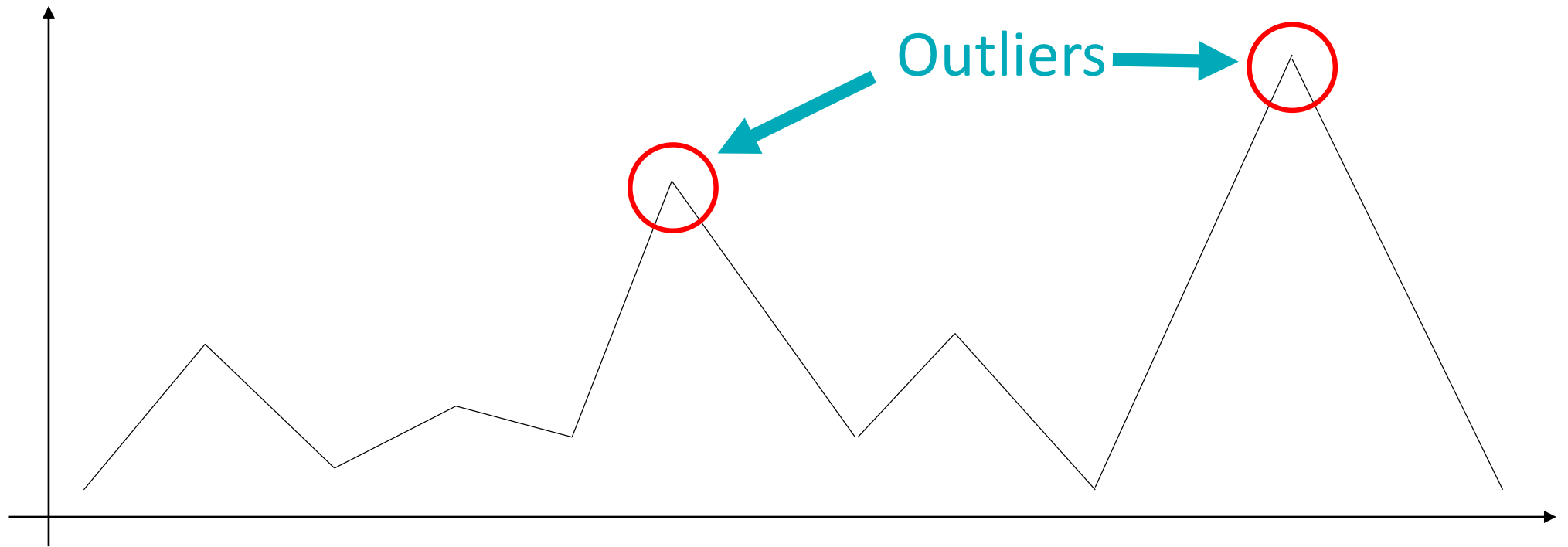
Machine Learning: Train and Test



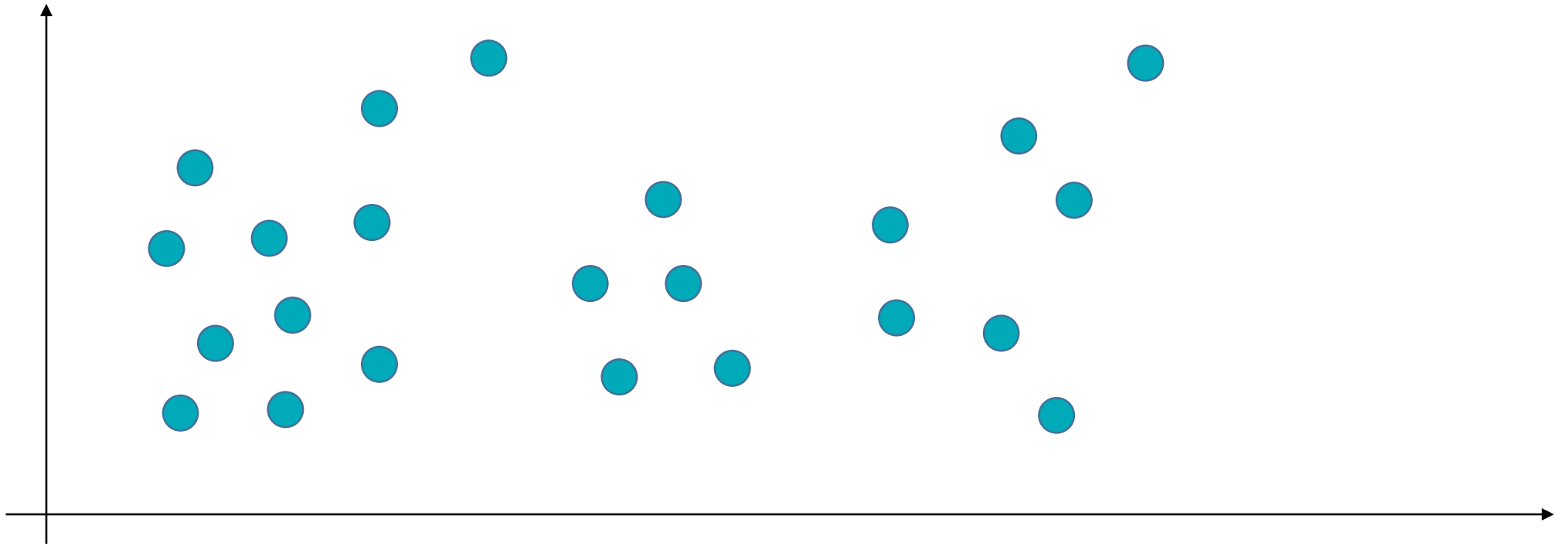
Machine Learning: Unsupervised



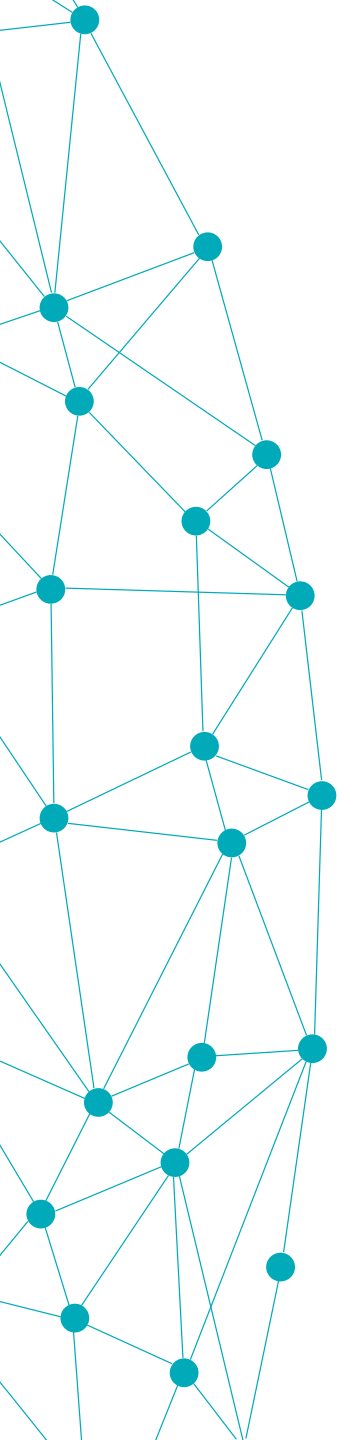
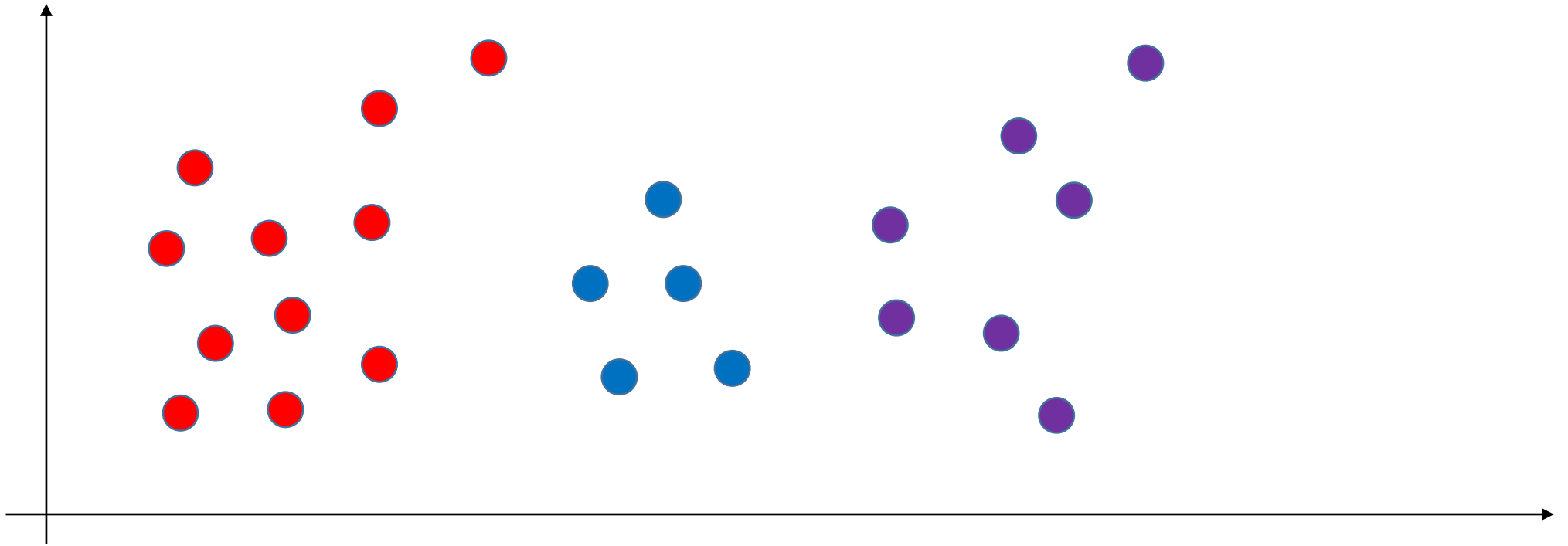
Machine Learning: Unsupervised



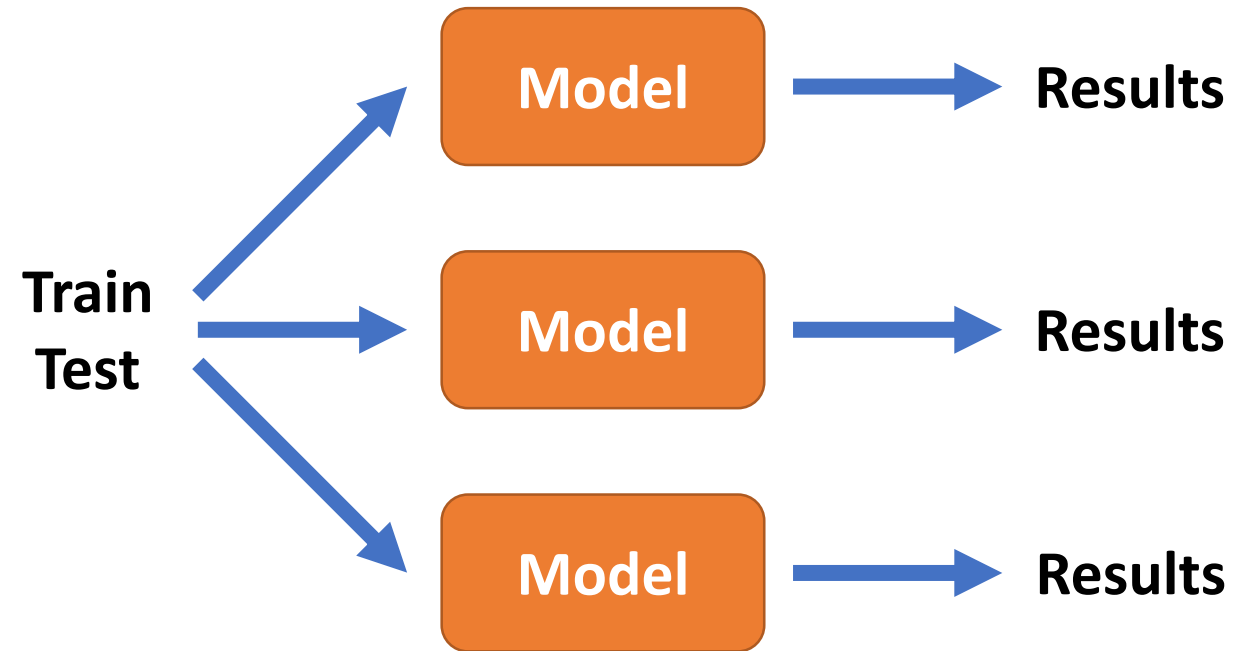
Machine Learning: Unsupervised



Machine Learning: Unsupervised



Machine Learning: Models





Classification



Machine Learning: Models

Supervised

Unsupervised

Classification

Clustering

Regression

Collaborative filtering

Machine Learning: Classification

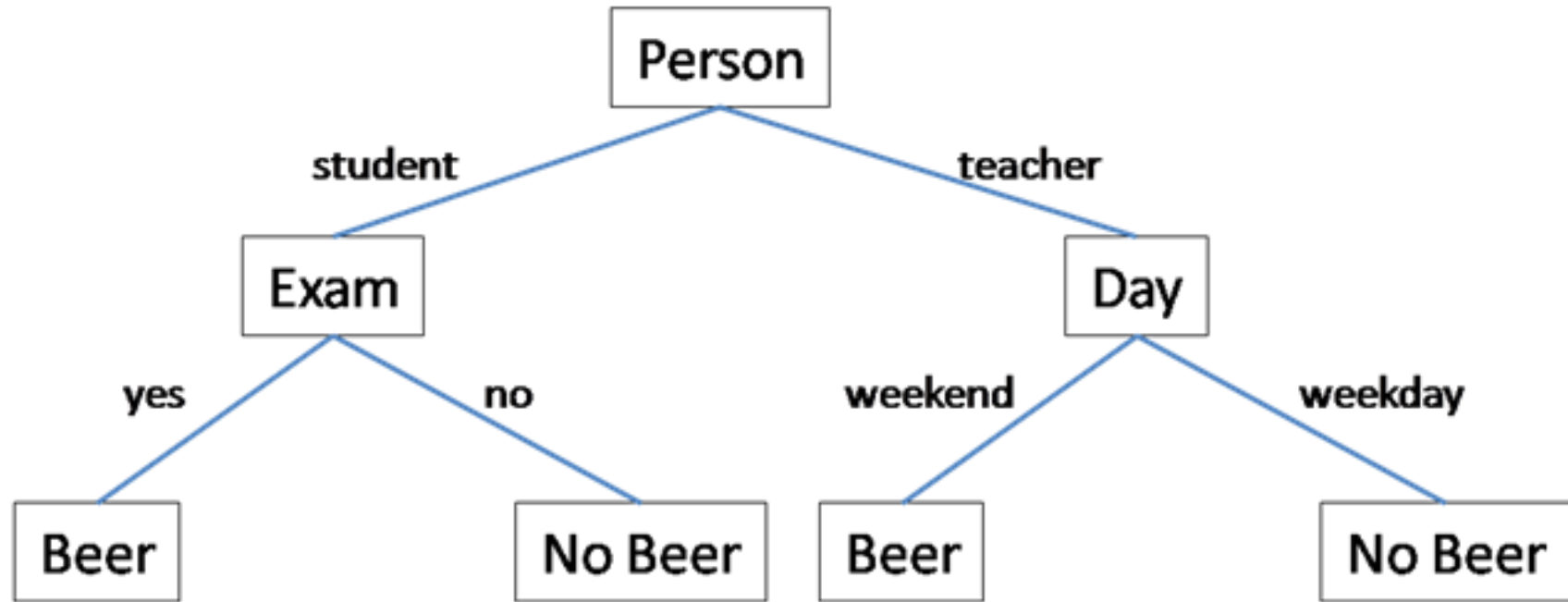
Man



Woman



Classification: Decision Tree

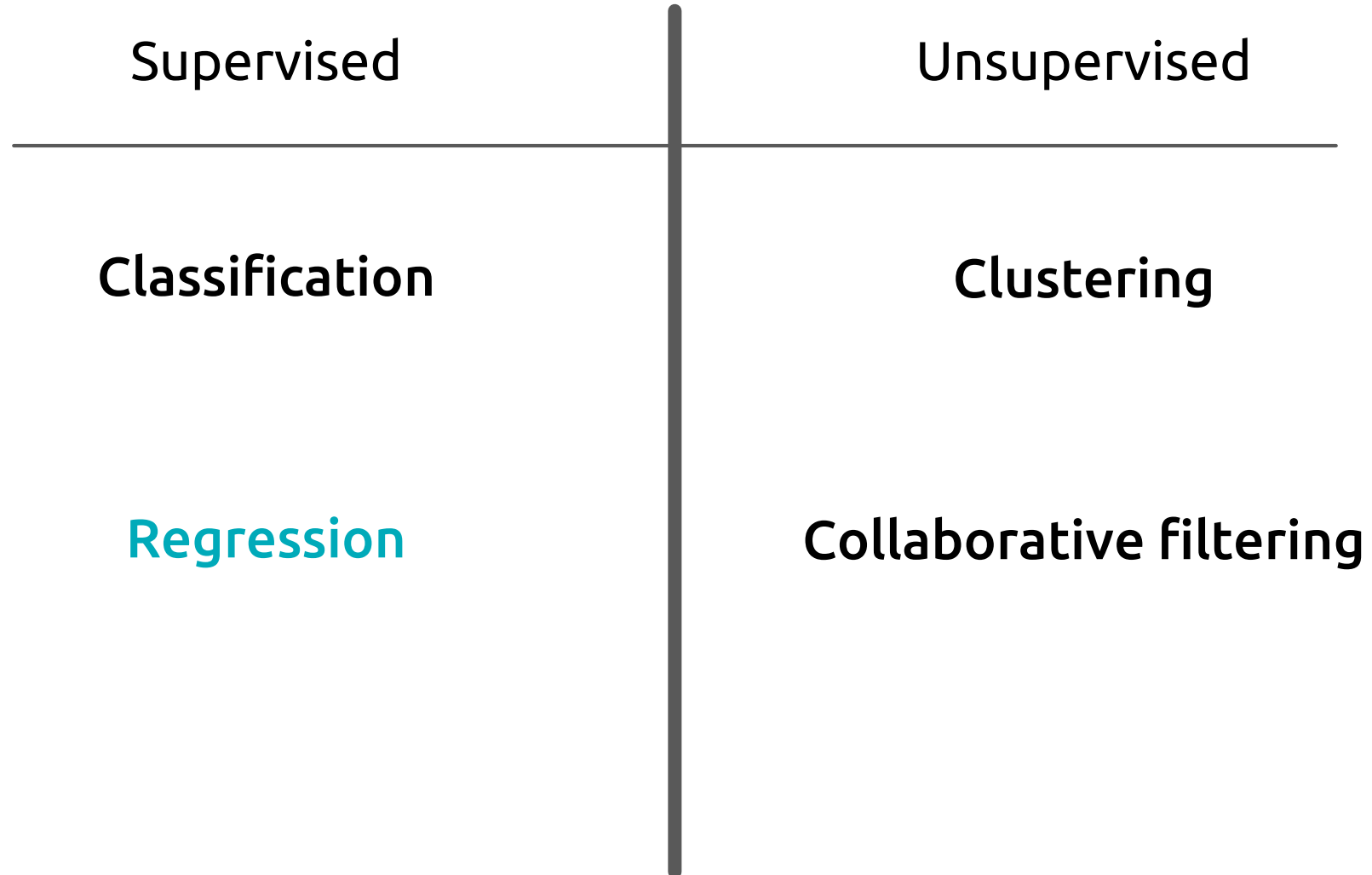


A decorative graphic on the left side of the slide, consisting of a network of light blue lines connecting small, solid blue circular nodes. The nodes are arranged in a somewhat vertical, irregular pattern, with some nodes having multiple connections, creating a web-like structure.

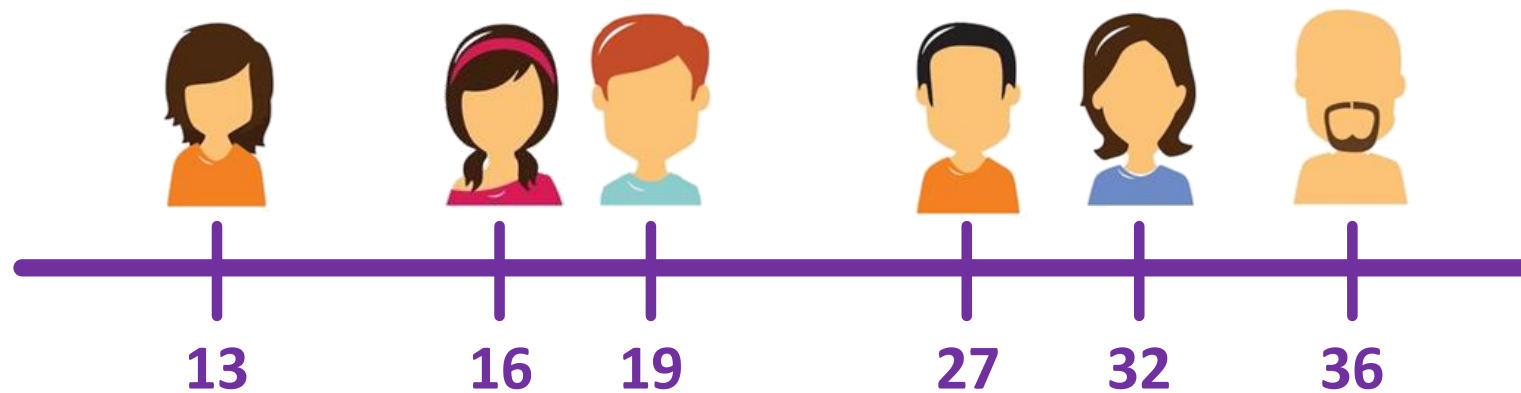
Regression



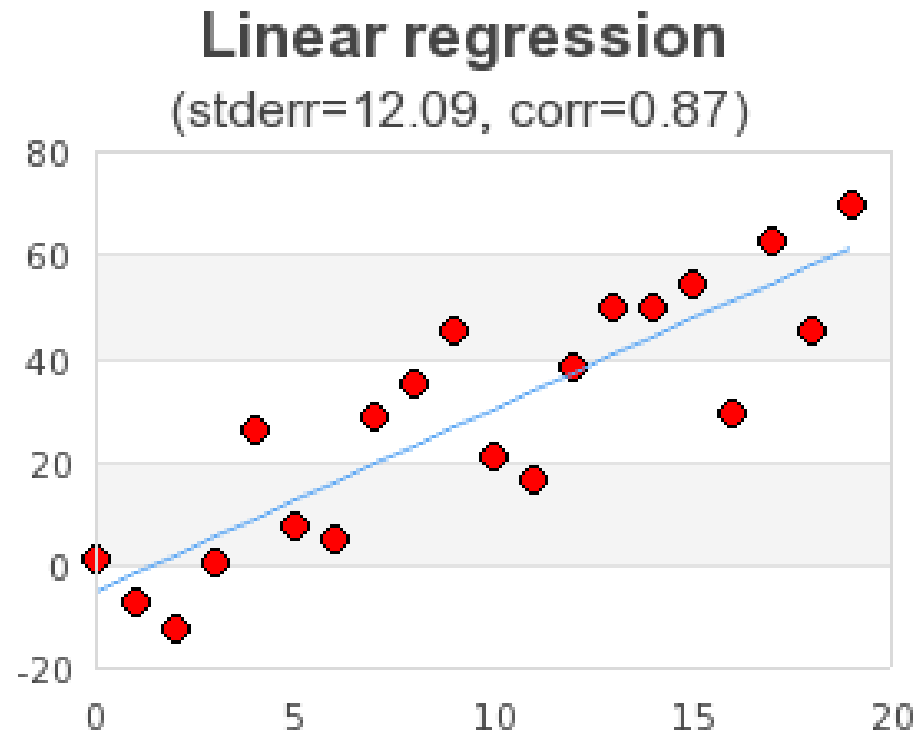
Machine Learning: Models



Machine Learning: Regression



Regression: Linear Regression

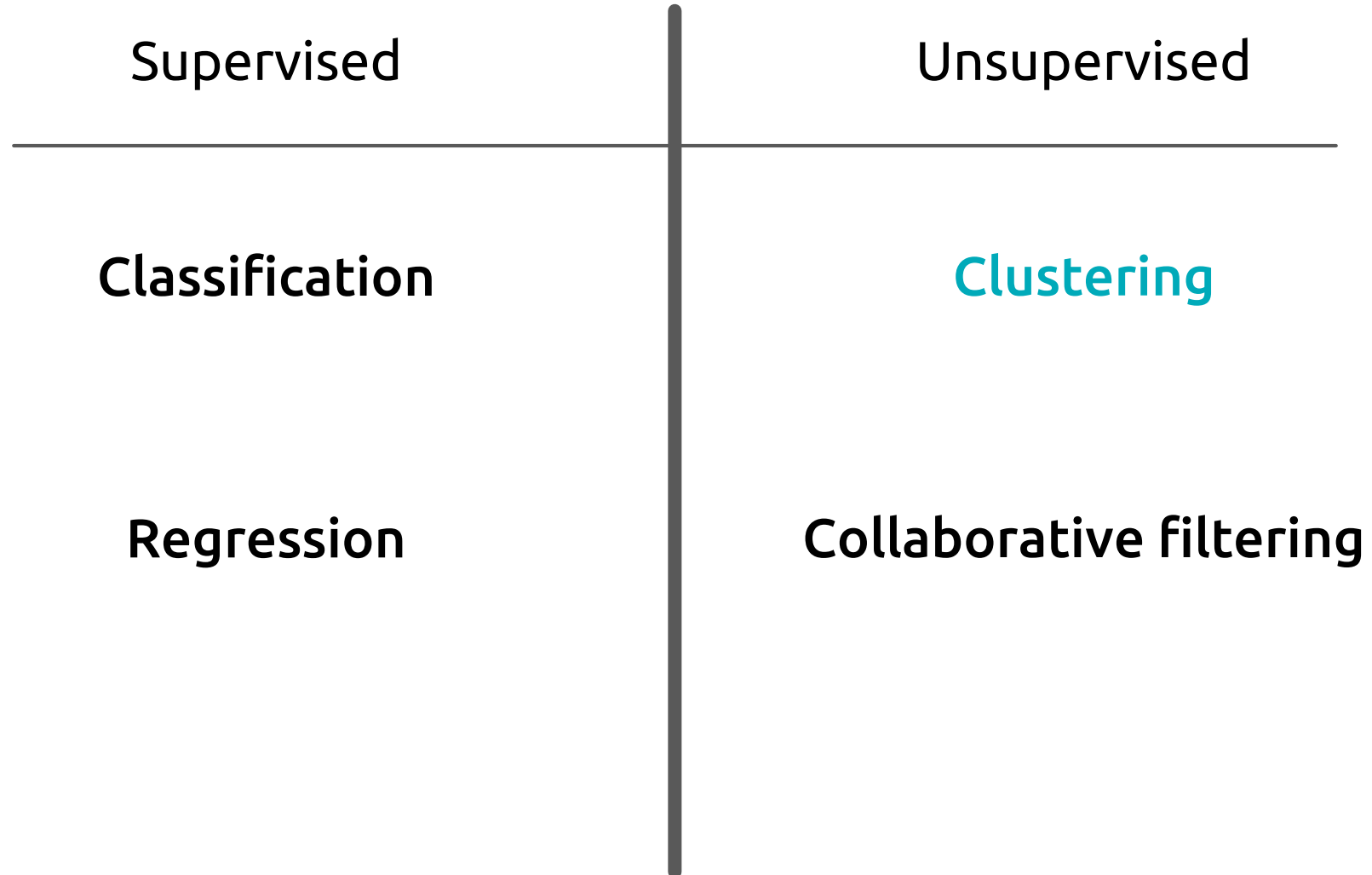


A network graph is positioned on the left side of the slide. It consists of approximately 20 blue circular nodes connected by thin, light blue lines. The nodes are arranged in a somewhat vertical, irregular pattern, with some nodes having multiple connections, creating a web-like structure. The background is a solid dark teal color.

Clustering



Machine Learning: Models



Machine Learning: Clustering

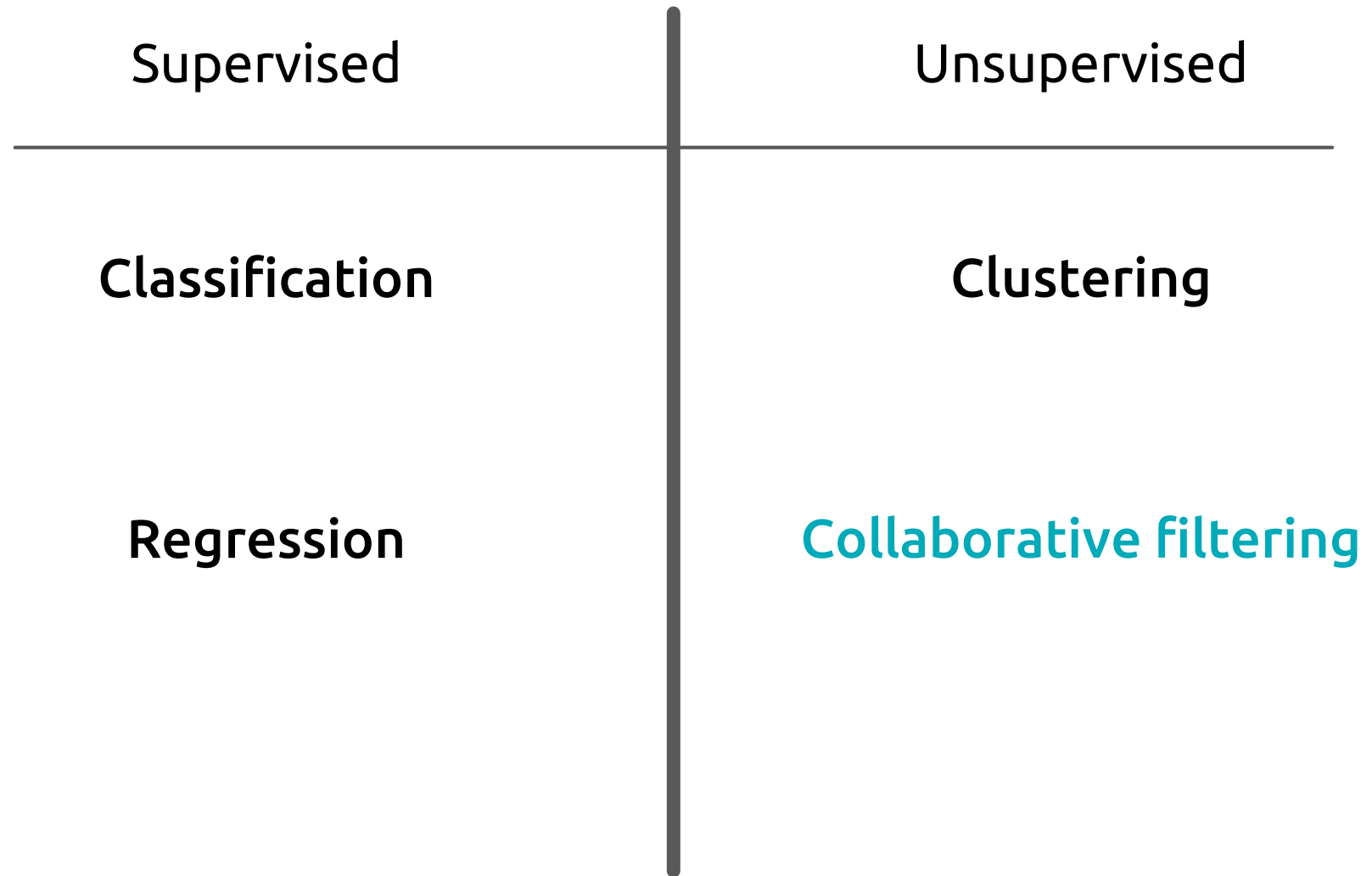




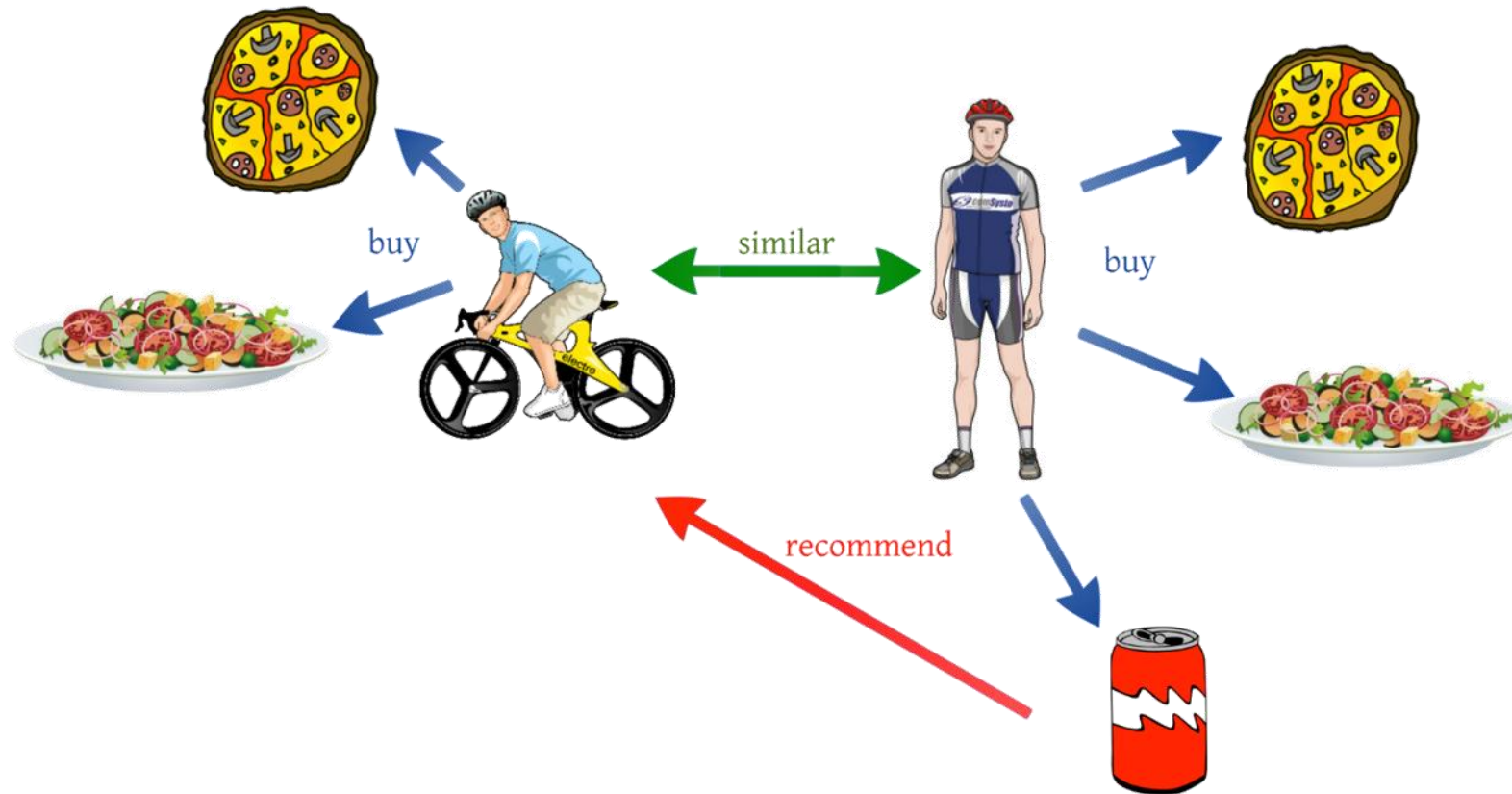
Collaborative filtering

A decorative network graph on the left side of the slide, consisting of blue circular nodes connected by thin blue lines, forming a complex web-like structure.

Machine learning: Models



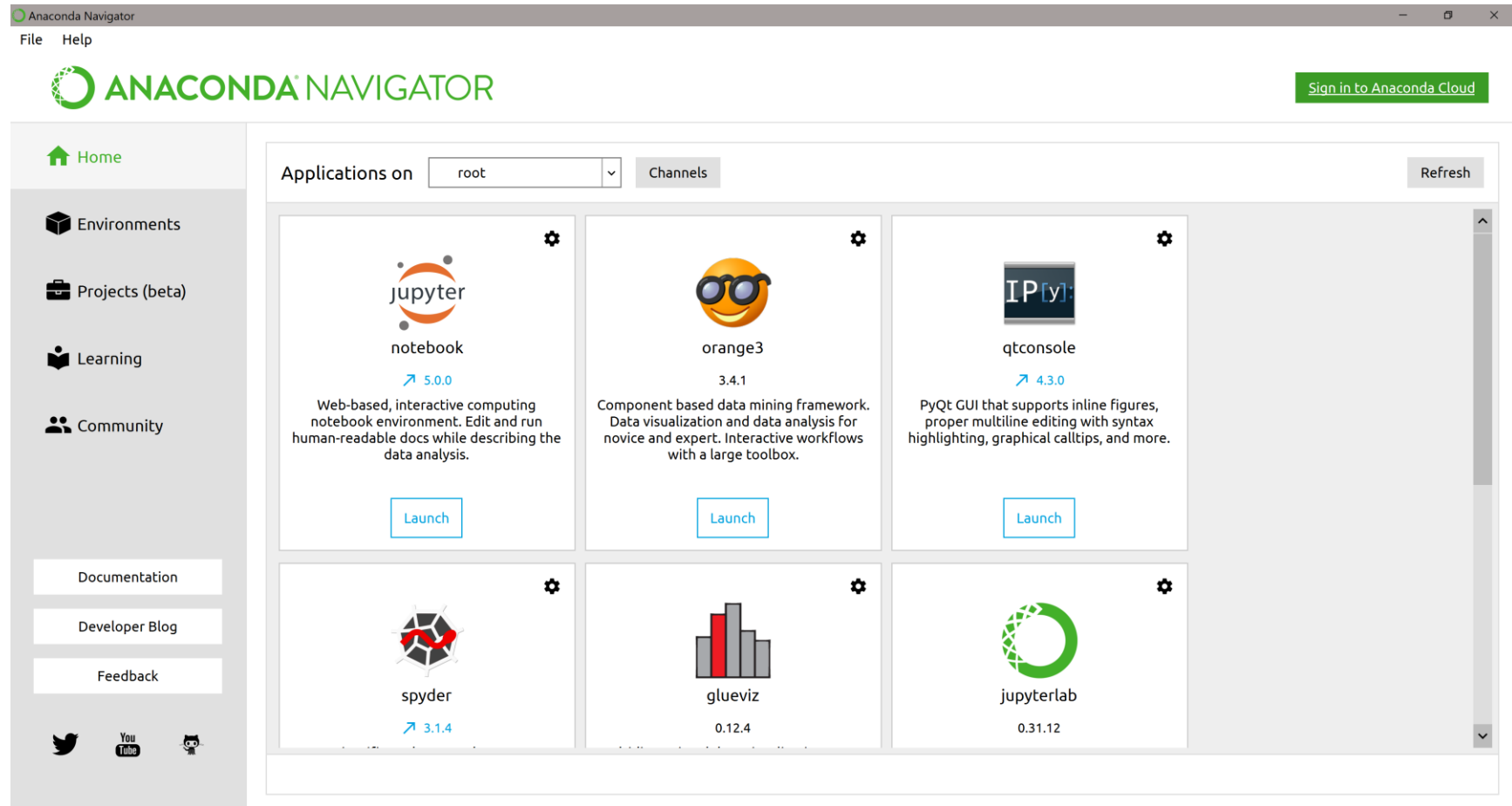
Filtering: User-based filtering



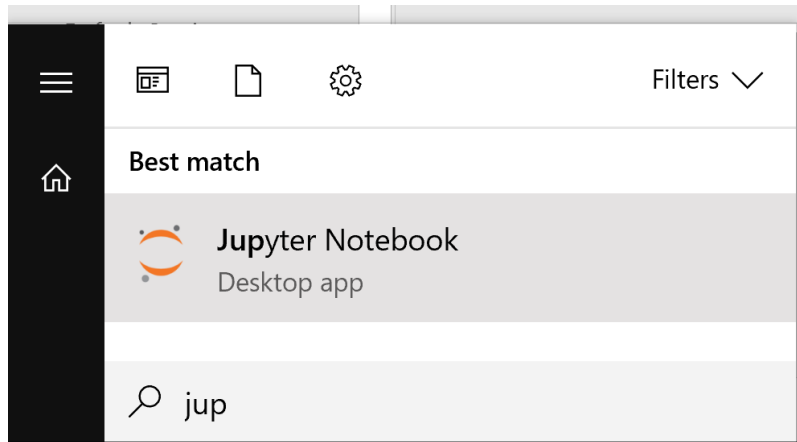
A decorative network graph is positioned on the left side of the slide. It consists of approximately 20 light blue circular nodes connected by thin, light blue lines, forming a complex, web-like structure that tapers towards the left edge.

Jupyter Notebook

Installing Jupyter notebook

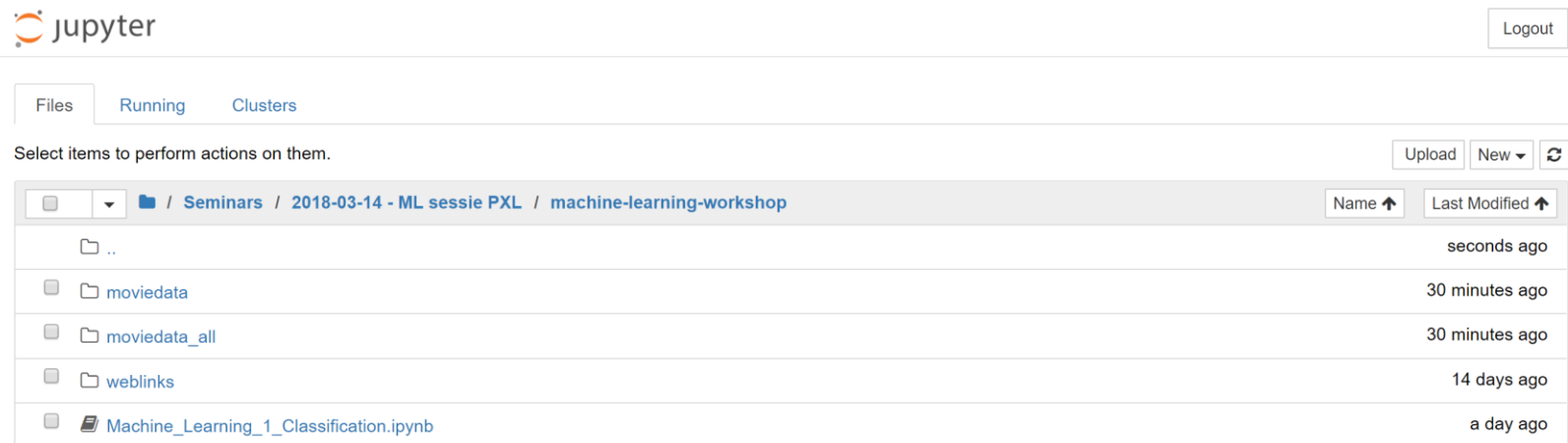


Starting Jupyter notebooks



```
C:\> Command Prompt
Microsoft Windows [Version 10.0.15063]
(c) 2017 Microsoft Corporation. All rights reserved.

C:\Users\oolsana>jupyter notebook
```





Exercise 1: Classification

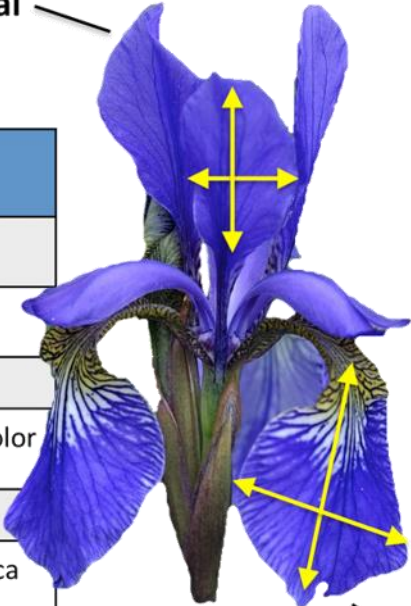
Classification: Iris dataset

Samples
(instances, observations)

	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
...					
50	6.4	3.5	4.5	1.2	Versicolor
...					
150	5.9	3.0	5.0	1.8	Virginica

Features
(attributes, measurements, dimensions)

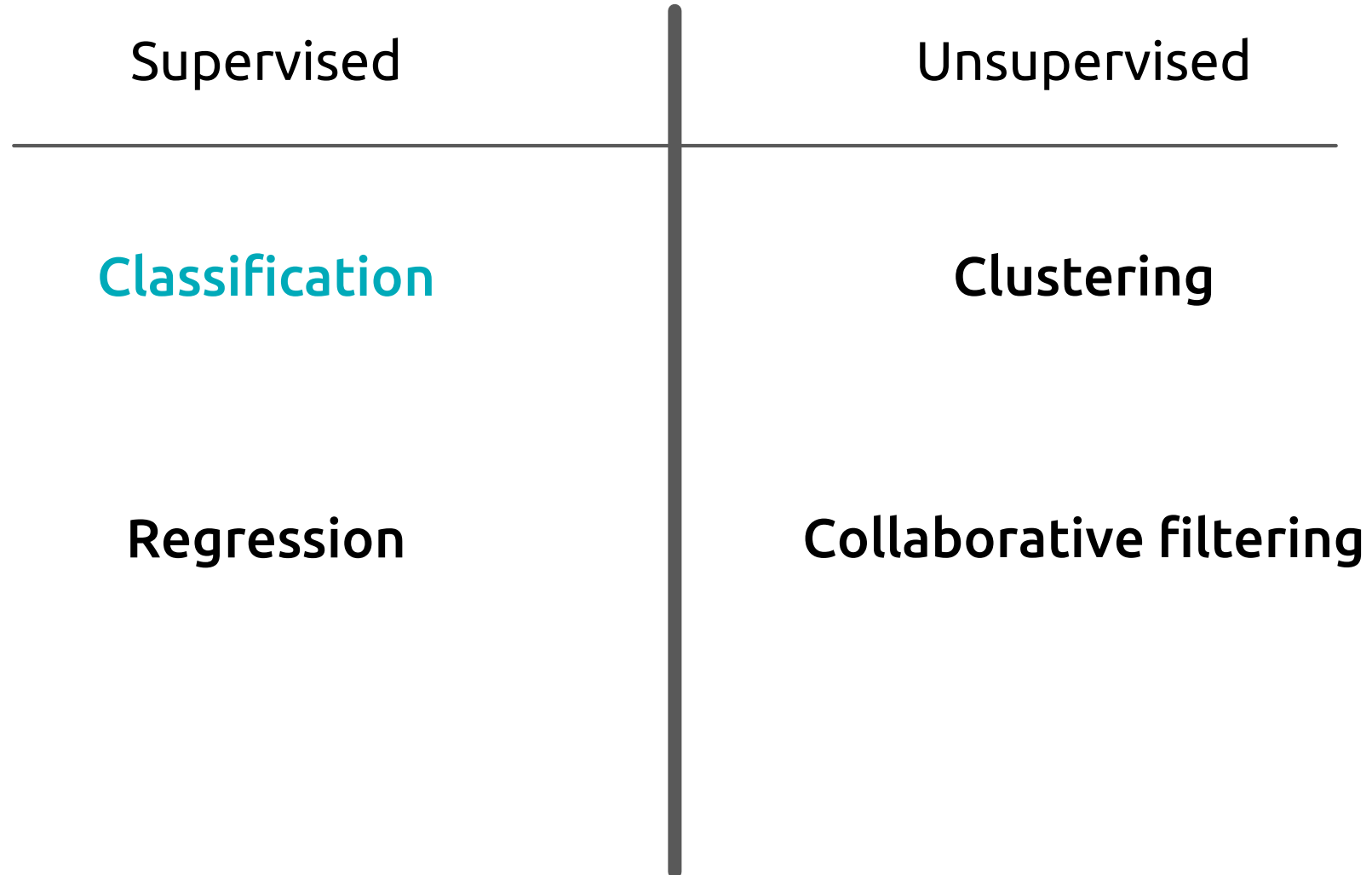
Class labels
(targets)



The diagram shows a blue Iris flower. A label 'Petal' points to the upper petals, and a label 'Sepal' points to the lower sepals. Yellow double-headed arrows indicate measurements: one vertical arrow on a petal and one diagonal arrow on a sepal.



Machine Learning: Models



A decorative network graph is located on the left side of the slide. It consists of approximately 20 light blue circular nodes connected by thin, light blue lines, forming a complex, interconnected web-like structure.

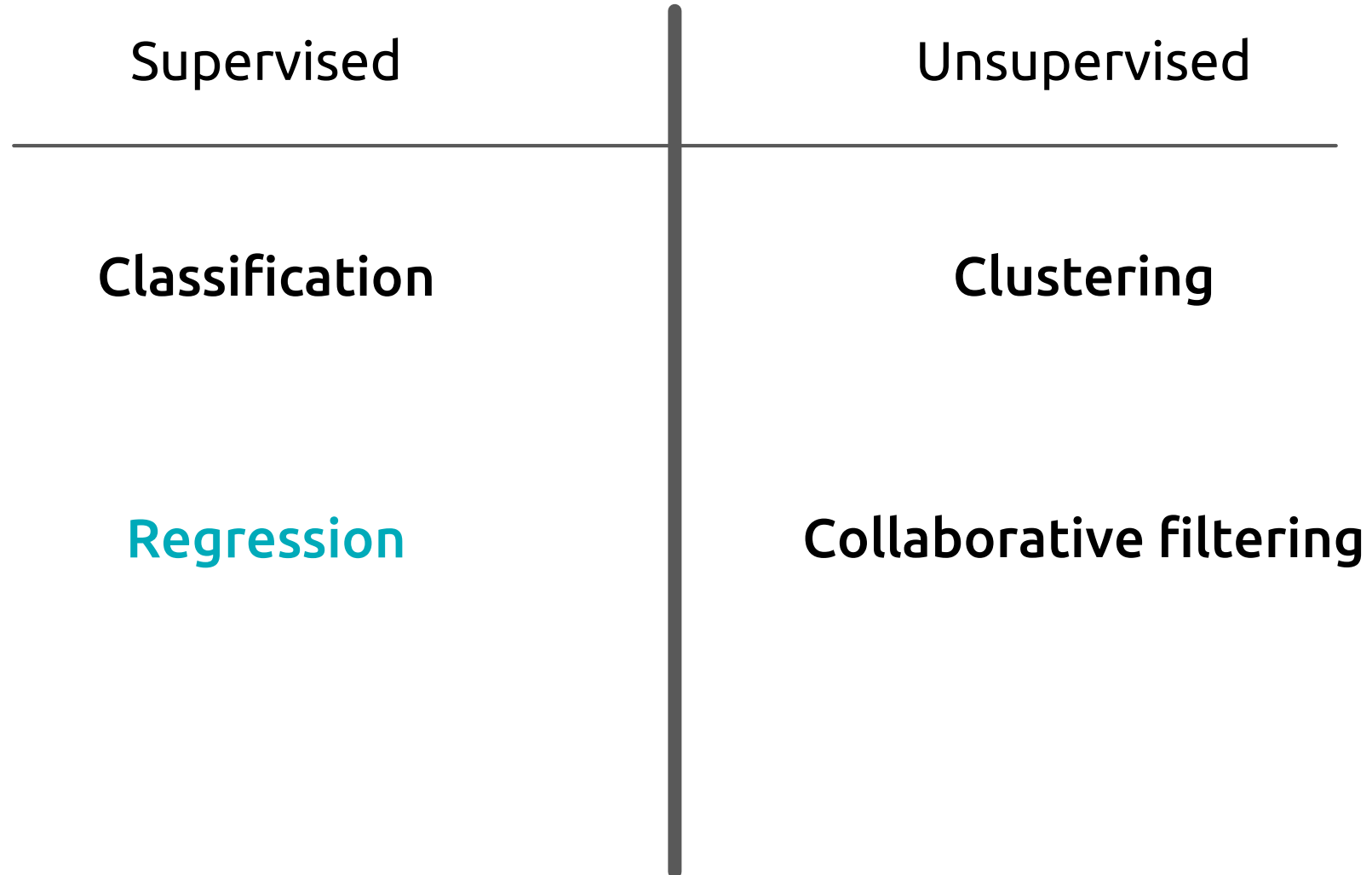
Exercise 2: Regression

Regression: Boston dataset

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	15.3	396.90	4.98
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	17.8	396.90	9.14
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	17.8	392.83	4.03
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	18.7	394.63	2.94
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	18.7	396.90	5.33
5	0.02985	0.0	2.18	0.0	0.458	6.430	58.7	6.0622	3.0	222.0	18.7	394.12	5.21
6	0.08829	12.5	7.87	0.0	0.524	6.012	66.6	5.5605	5.0	311.0	15.2	395.60	12.43
7	0.14455	12.5	7.87	0.0	0.524	6.172	96.1	5.9505	5.0	311.0	15.2	396.90	19.15
8	0.21124	12.5	7.87	0.0	0.524	5.631	100.0	6.0821	5.0	311.0	15.2	386.63	29.93
9	0.17004	12.5	7.87	0.0	0.524	6.004	85.9	6.5921	5.0	311.0	15.2	386.71	17.10



Machine Learning: Models





Exercise 3: Collaborative filtering

Filtering: movie reviews



Filtering: movie reviews



A decorative network graph on the left side of the slide, consisting of blue circular nodes connected by thin blue lines, forming a complex web-like structure.

Machine learning: Models

