PA214 - Visualization II

Jan Byška

Machine Learning Overview

HCI LAB

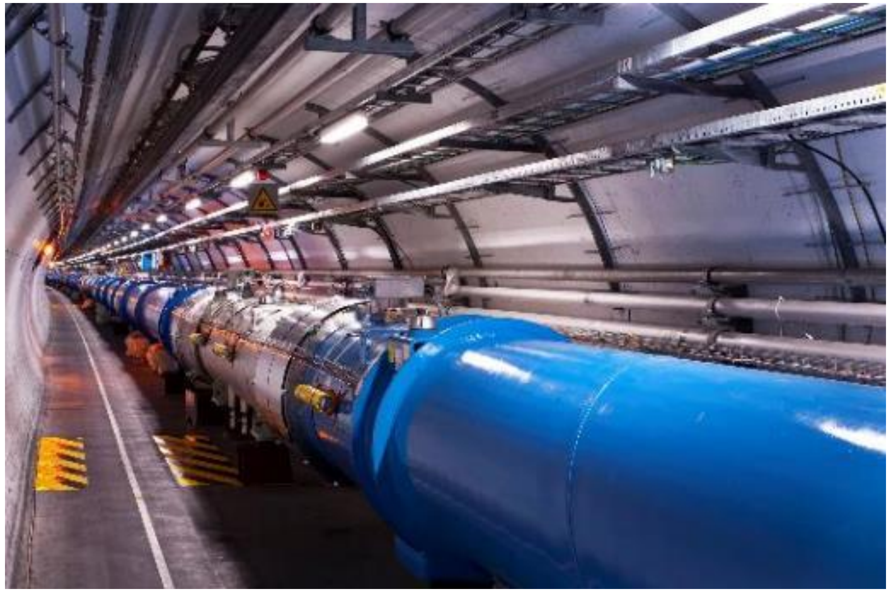
visitlab

Challenges with Data (4V's of Data)

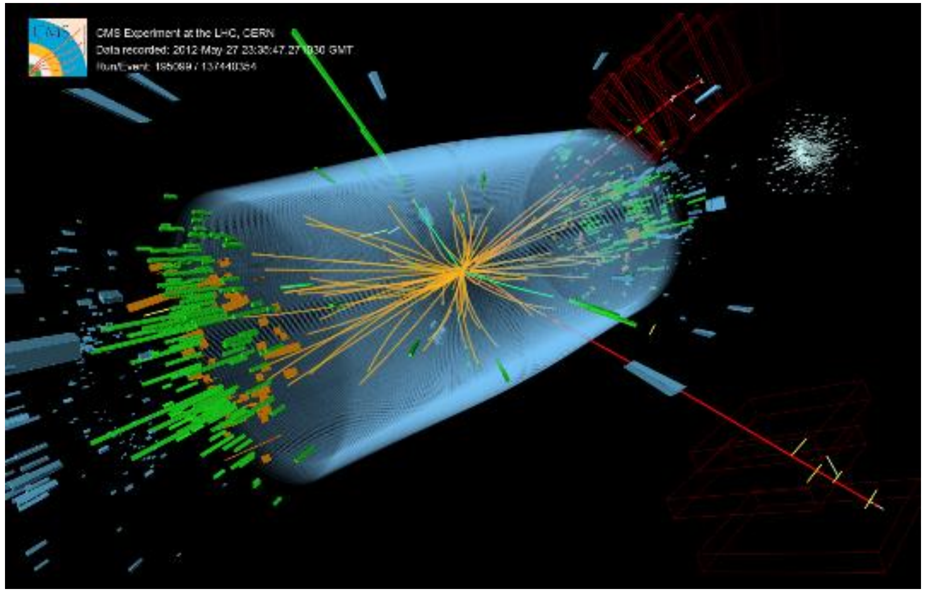
- **Volume:** Dealing with large volumes of data.

Challenges with Data (4V’s of Data)

- **Volume:** Dealing with large volumes of data.
- **Velocity:** Handling the speed at which data is generated.



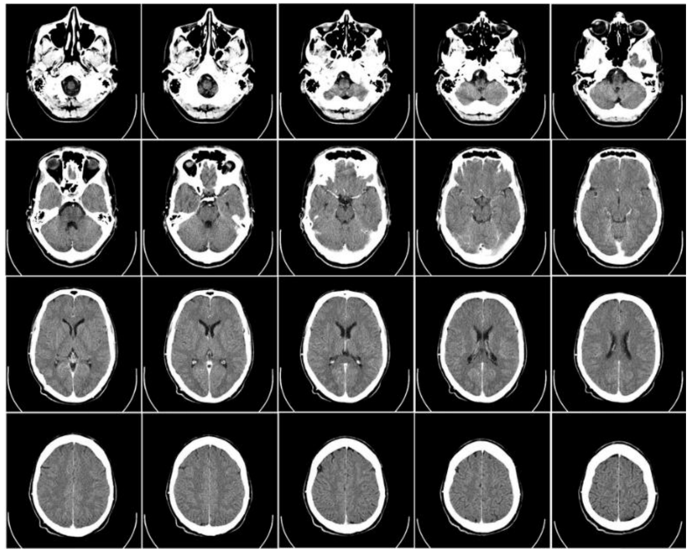
By Maximilien Brice (CERN) – Wikimedia Commons, CC BY-SA 3.0

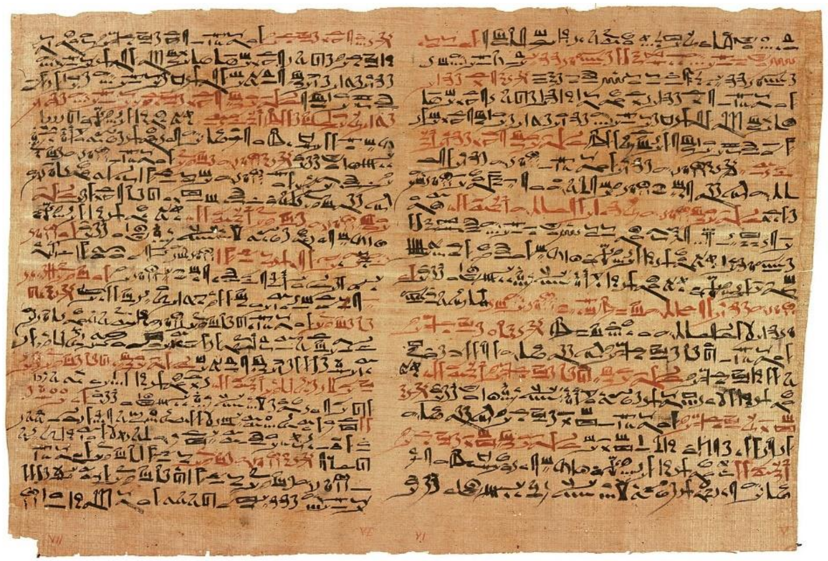

CMS Experiment at the LHC, CERN
Data recorded: 2012-May-27 23:35:47.271030 GMT
Run/Event: 195099 / 137440354

By Thomas Mc Cauley; Lucas Taylor - CMS Collaboration, CC BY-SA 4.0

Challenges with Data (4V’s of Data)

- **Volume:** Dealing with large volumes of data.
- **Velocity:** Handling the speed at which data is generated.
- **Variety:** Managing different types of data (structured, unstructured, semi-structured).
- **Veracity:** Ensuring data quality and reliability.

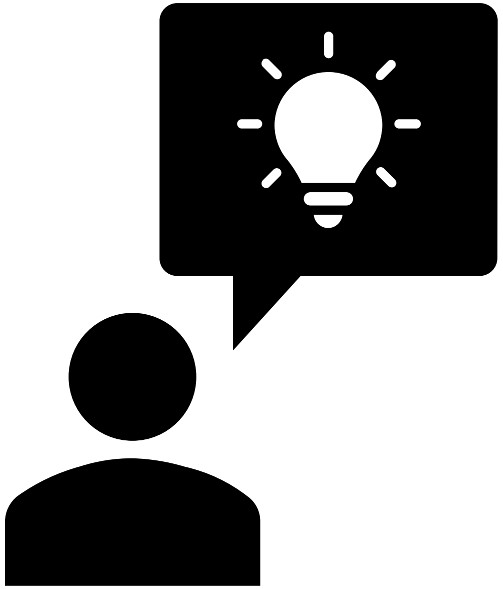




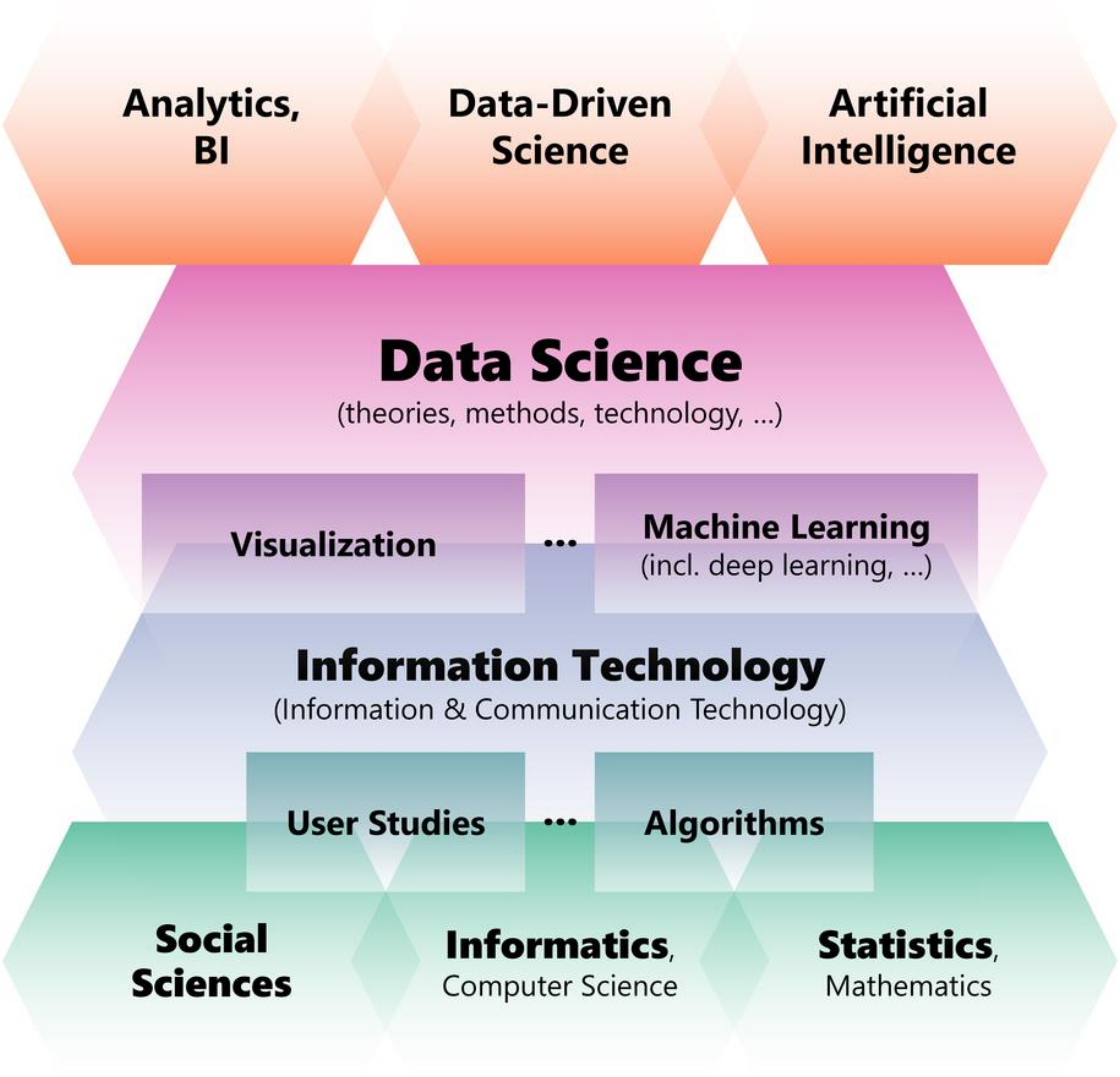
Challenges with Data (5V's of Data)

- **Volume:** Dealing with large volumes of data.
- **Velocity:** Handling the speed at which data is generated.
- **Variety:** Managing different types of data (structured, unstructured, semi-structured).
- **Veracity:** Ensuring data quality and reliability.

- **Value:** Extracting meaningful insights from data.



(Visual) Data Science



Analytics, BI

Data-Driven Science

Artificial Intelligence

Data Science

(theories, methods, technology, ...)

Visualization ... **Machine Learning** (incl. deep learning, ...)

Information Technology

(Information & Communication Technology)

User Studies ... **Algorithms**

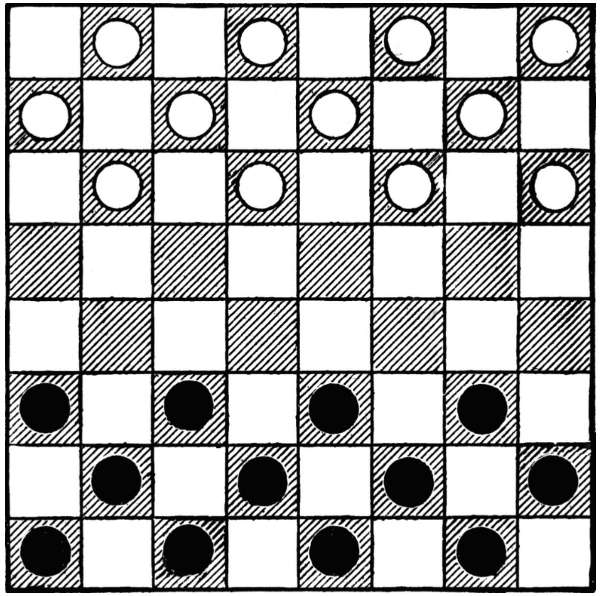
Social Sciences

Informatics, Computer Science

Statistics, Mathematics

Introduction

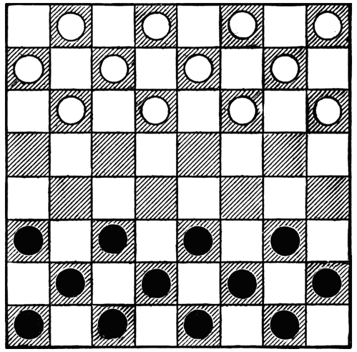
- Machine Learning:
  - "Field of study that gives computers the ability to learn without being explicitly programmed." (1959) Arthur Samuel (pioneer in AI & ML)



Introduction

Why ML?

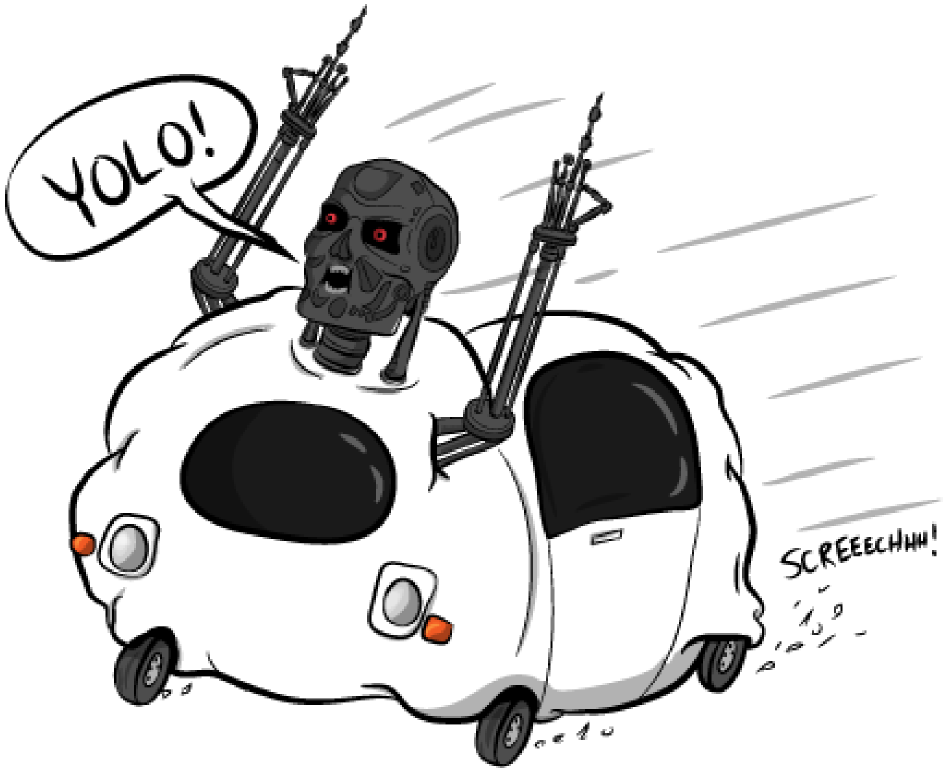
- Hard to write a computer program
  - 10^{40} legal moves variations (between 10^{111} and 10^{123})
  - Cannot be brute forced
  - Cannot be modeled
  - **Cannot be visualized**
- Best players
  - Rely on experience
- Computers
  - Can obtain „experience“ much faster



Fabiano Caruana (Photo by: Soeren Stache)

Application Examples

Autonomous cars/drones



An abandoned prototype.

Source: http://theoatmeal.com/blog/google_self_driving_car

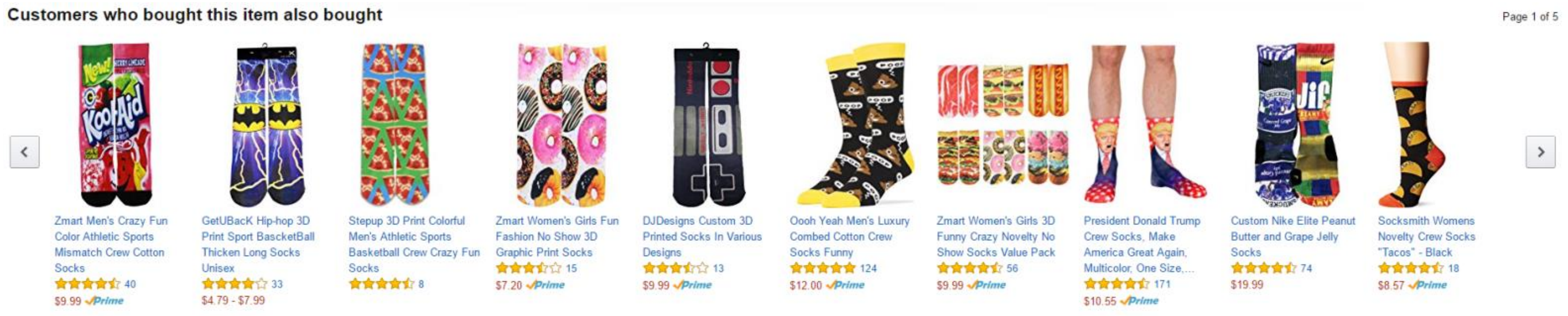
Application Examples

Adaption in games, imitation learning

- Black & White 3 (2001)
  - Avatar learns from the player

Application Examples

- The main problems solved by ML
  - classification
  - clustering
  - dimensionality reduction, embedding
  - outlier detection
  - prediction
  - …

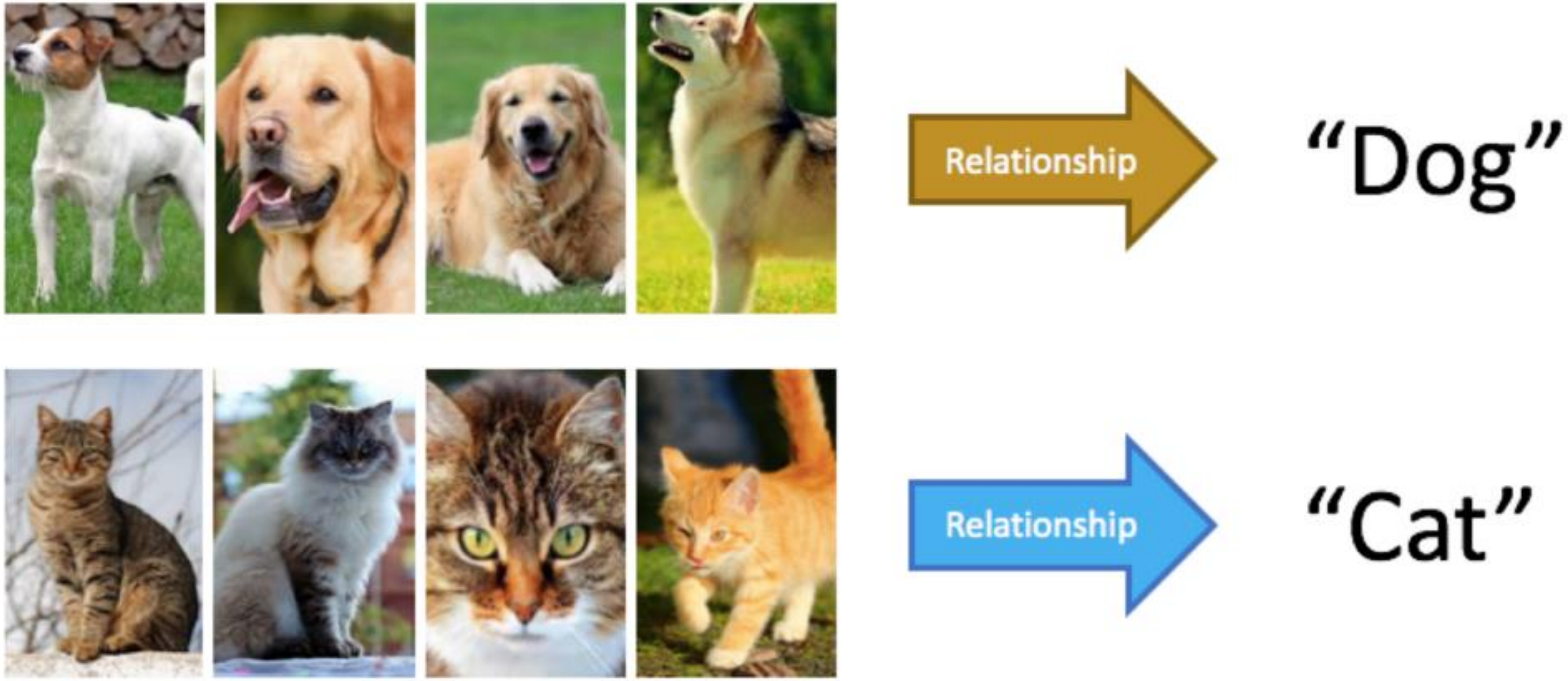


Customers who bought this item also bought

Page 1 of 5

- Zmart Men's Crazy Fun Color Athletic Sports Mismatch Crew Cotton Socks ★★★★☆ 40 $9.99 Prime
- GetUBacK Hip-hop 3D Print Sport BascketBall Thicken Long Socks Unisex ★★★★☆ 33 $4.79 - $7.99
- Stepup 3D Print Colorful Men's Athletic Sports Basketball Crew Crazy Fun Socks ★★★★☆ 8
- Zmart Women's Girls Fun Fashion No Show 3D Graphic Print Socks ★★★☆☆ 15 $7.20 Prime
- DJDesigns Custom 3D Printed Socks In Various Designs ★★★☆☆ 13 $9.99 Prime
- Oooh Yeah Men's Luxury Combed Cotton Crew Socks Funny ★★★★★ 124 $12.00 Prime
- Zmart Women's Girls 3D Funny Crazy Novelty No Show Socks Value Pack ★★★★☆ 56 $9.99 Prime
- President Donald Trump Crew Socks, Make America Great Again, Multicolor, One Size,... ★★★★☆ 171 $10.55 Prime
- Custom Nike Elite Peanut Butter and Grape Jelly Socks ★★★★☆ 74 $19.99
- Socksmith Womens Novelty Crew Socks "Tacos" - Black ★★★★☆ 18 $8.57 Prime

Application Examples



https://www.edureka.co/blog/how-to-become-a-machine-learning-engineer/

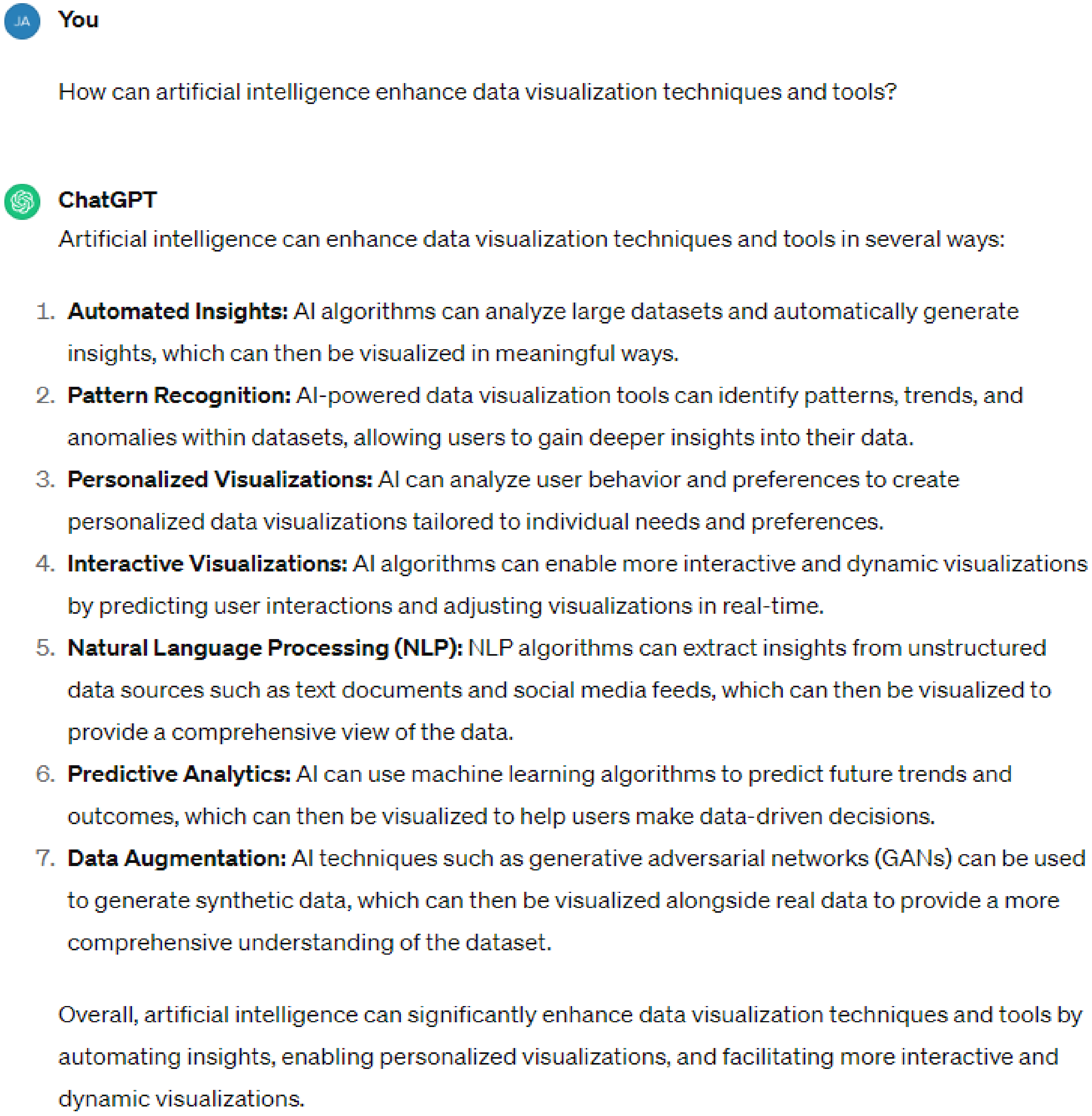
Application Examples

ChatGPT

GitHub Copilot

Application Examples

- ChatGPT



You

How can artificial intelligence enhance data visualization techniques and tools?

ChatGPT

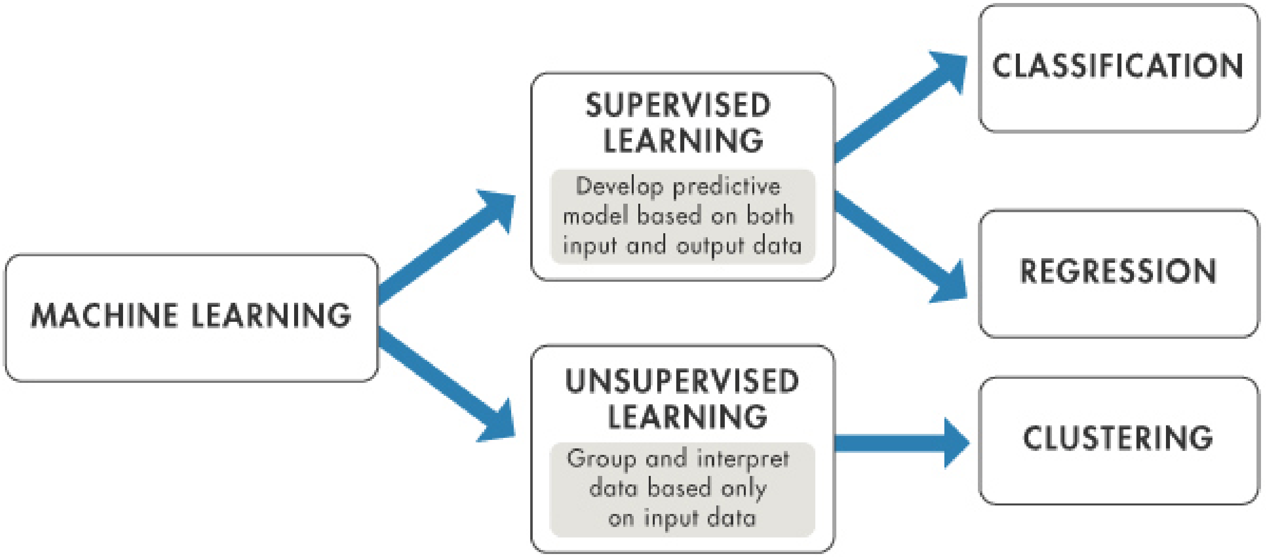
Artificial intelligence can enhance data visualization techniques and tools in several ways:

1. **Automated Insights:** AI algorithms can analyze large datasets and automatically generate insights, which can then be visualized in meaningful ways.
2. **Pattern Recognition:** AI-powered data visualization tools can identify patterns, trends, and anomalies within datasets, allowing users to gain deeper insights into their data.
3. **Personalized Visualizations:** AI can analyze user behavior and preferences to create personalized data visualizations tailored to individual needs and preferences.
4. **Interactive Visualizations:** AI algorithms can enable more interactive and dynamic visualizations by predicting user interactions and adjusting visualizations in real-time.
5. **Natural Language Processing (NLP):** NLP algorithms can extract insights from unstructured data sources such as text documents and social media feeds, which can then be visualized to provide a comprehensive view of the data.
6. **Predictive Analytics:** AI can use machine learning algorithms to predict future trends and outcomes, which can then be visualized to help users make data-driven decisions.
7. **Data Augmentation:** AI techniques such as generative adversarial networks (GANs) can be used to generate synthetic data, which can then be visualized alongside real data to provide a more comprehensive understanding of the dataset.

Overall, artificial intelligence can significantly enhance data visualization techniques and tools by automating insights, enabling personalized visualizations, and facilitating more interactive and dynamic visualizations.

Principles

- Machine learning:
  - Data «tells» what the «good answers» are (training).
  - No explicit commands coded
  - Key point of ML is the training of the algorithm
  - Three main learning styles:
    - supervised
    - unsupervised
    - semi-supervised



MACHINE LEARNING

SUPERVISED LEARNING

Develop predictive model based on both input and output data

UNSUPERVISED LEARNING

Group and interpret data based only on input data

CLASSIFICATION

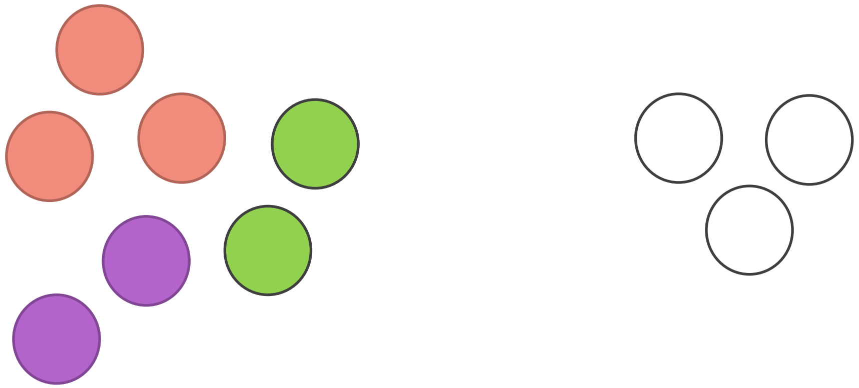
REGRESSION

CLUSTERING

https://www.mathworks.com

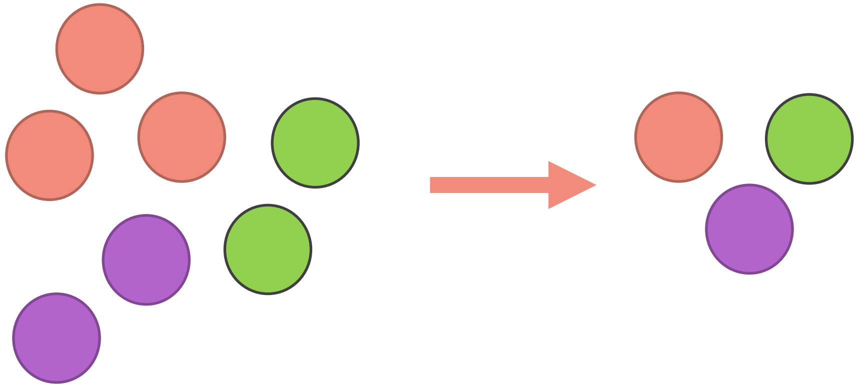
Learning Styles

- Supervised learning
    - Labeled input
    - Model prepared through training that requires predictions, corrected when wrong
    - Problem examples: classification, regression
    - Algorithmic examples: neural networks, Bayes classifiers, decisions trees...
- Unsupervised learning
- Semi-supervised learning

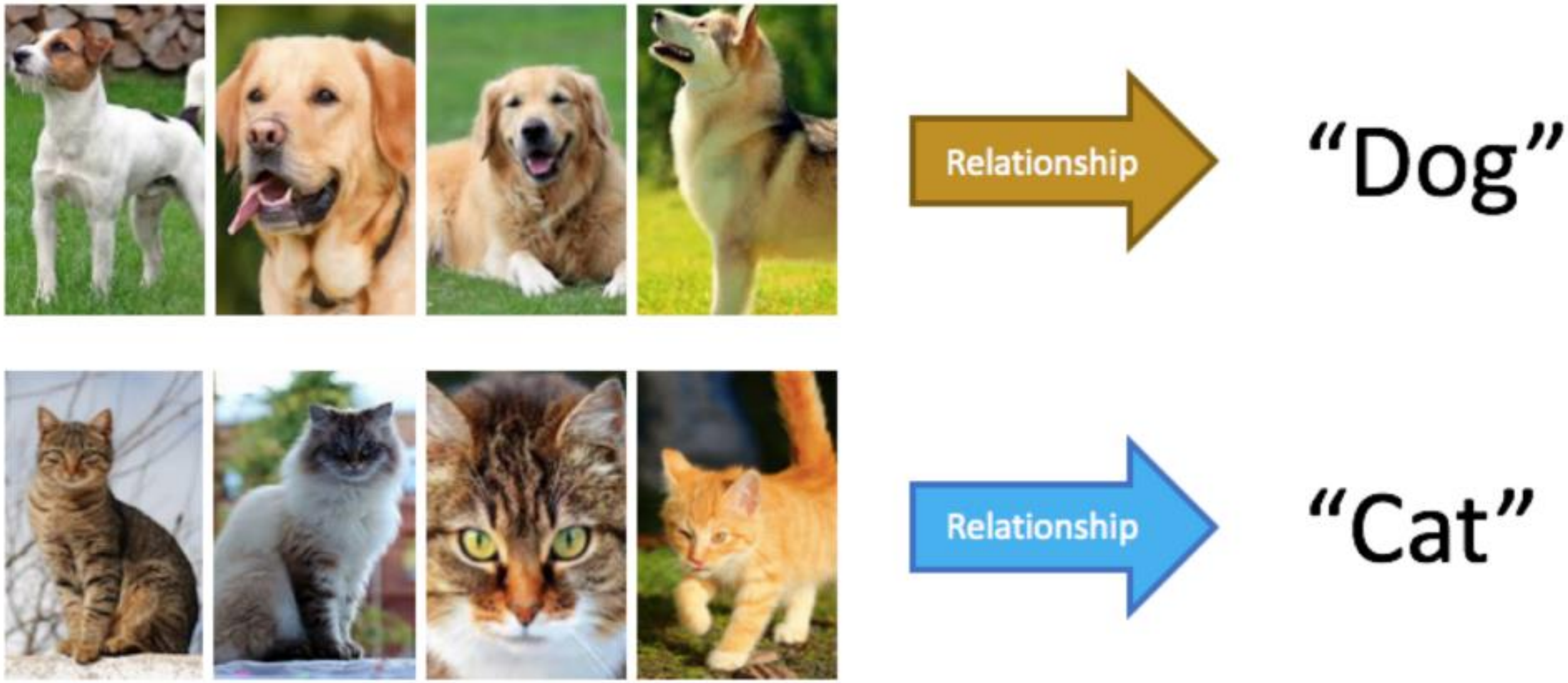


Learning Styles

- Supervised learning
    - Labeled input
    - Model prepared through training that requires predictions, corrected when wrong
    - Problem examples: classification, regression
    - Algorithmic examples: neural networks, Bayes classifiers, decisions trees…
- Unsupervised learning
- Semi-supervised learning



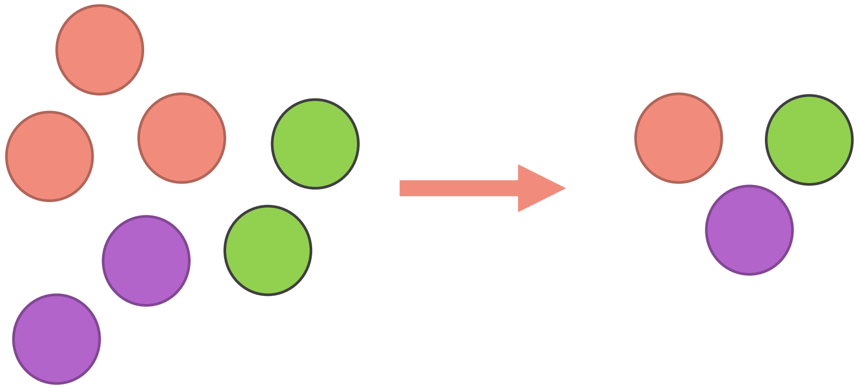
Supervised Learning



https://www.edureka.co/blog/how-to-become-a-machine-learning-engineer/

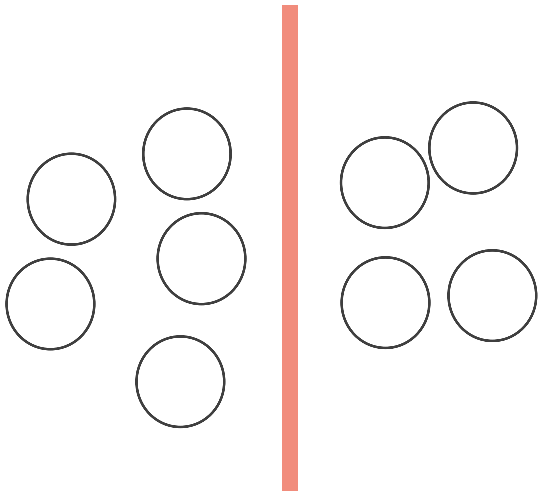
Learning Styles

- Supervised learning
  - Labeled input
  - Model prepared through training that requires predictions, corrected when wrong
  - Problem examples: classification, regression
  - Algorithmic examples: neural networks, Bayes classifiers, decisions trees...
- Unsupervised learning
- Semi-supervised learning



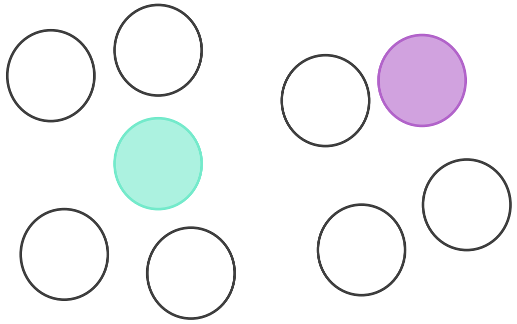
Learning Styles

- Supervised Learning
- Unsupervised Learning
  - Input not labeled, no known result
  - Model is prepared by deducing structures in the data
  - Problem examples: clustering, dimensionality reduction
  - Algorithmic examples: a priori algorithm, **k-means**
- Semi-Supervised Learning



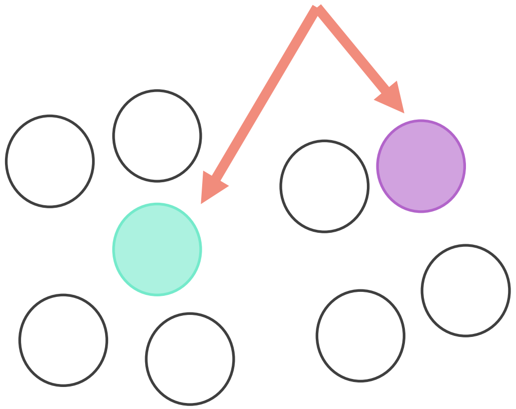
Learning Styles

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
  - Input is a mixture of labeled and unlabeled data
  - Model has to recognize structures and make predictions
  - Problem examples: classification, regression
  - Algorithmic examples: label propagation (adaptive learning)



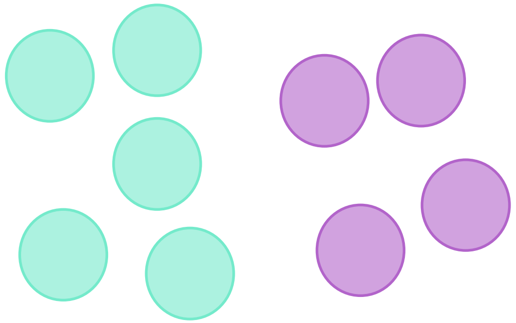
Learning Styles

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
  - Input is a mixture of labeled and unlabeled data
  - Model has to recognize structures and make predictions
  - Problem examples: classification, regression
  - Algorithmic examples: label propagation (adaptive learning)



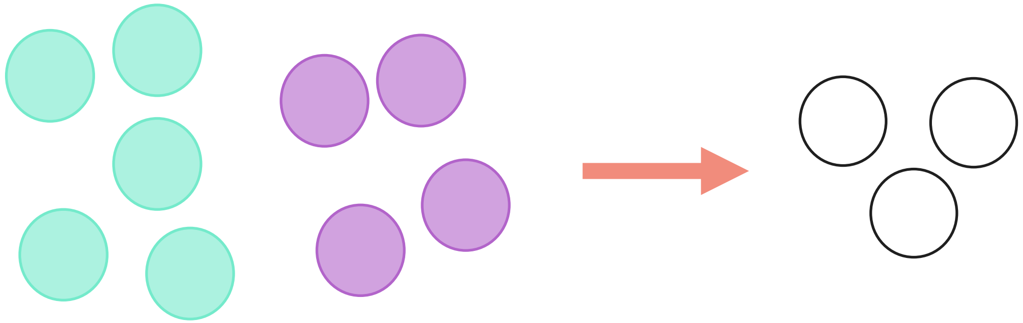
Learning Styles

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
  - Input is a mixture of labeled and unlabeled data
  - Model has to recognize structures and make predictions
  - Problem examples: classification, regression
  - Algorithmic examples: label propagation (adaptive learning)



Learning Styles

- Supervised Learning
- Unsupervised Learning
- Semi-Supervised Learning
    - Input is a mixture of labeled and unlabeled data
    - Model has to recognize structures and make predictions
    - Problem examples: classification, regression
    - Algorithmic examples: label propagation (adaptive learning)



Principles

- The success of a ML algorithm is highly dependent on two key decisions:
  - Data representation
  - Choice of the classifier

Principles

- The success of a ML algorithm is highly dependent on two key decisions:
    - Data representation
    - Choice of the classifier

- Data representation:
    - What is the important information in the data?
    - How simple can the data be represented?
    - Is a basis change needed for a better representation?

Principles

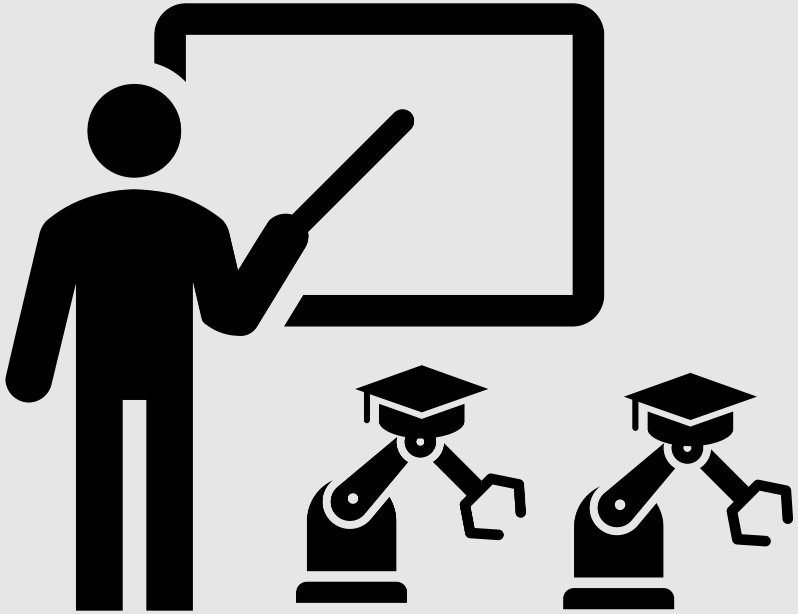
- The success of a ML algorithm is highly dependent on two key decisions:
  - Data representation
  - Choice of the classifier

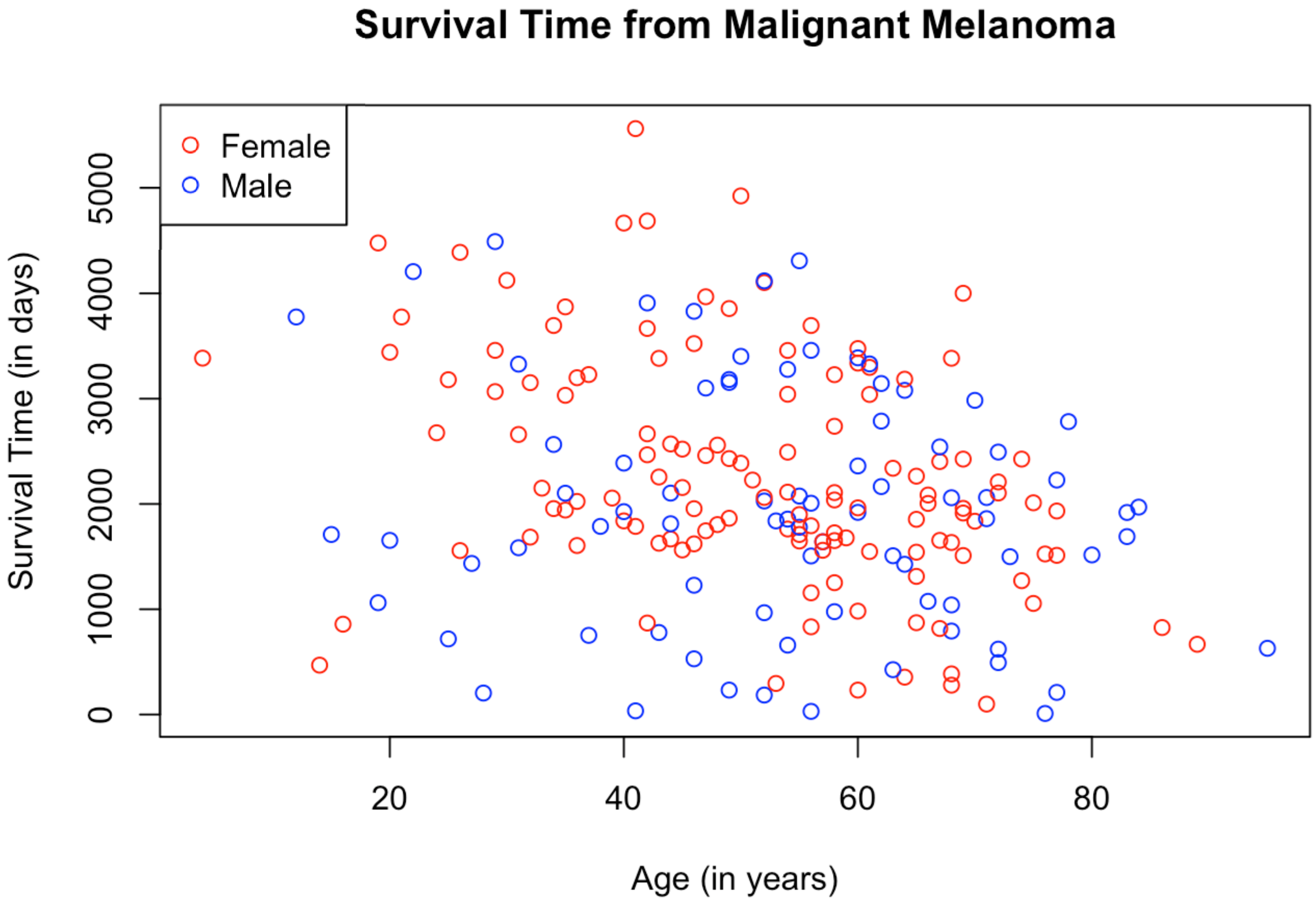
- Classifier choice:
  - The classifier makes the «decisions», a badly suited classifier will make bad decisions
  - The choice of the classifier is dependent on:
    - Size of the data
    - Variance of the data
    - Bias of the data
    - ....

Algorithm Families

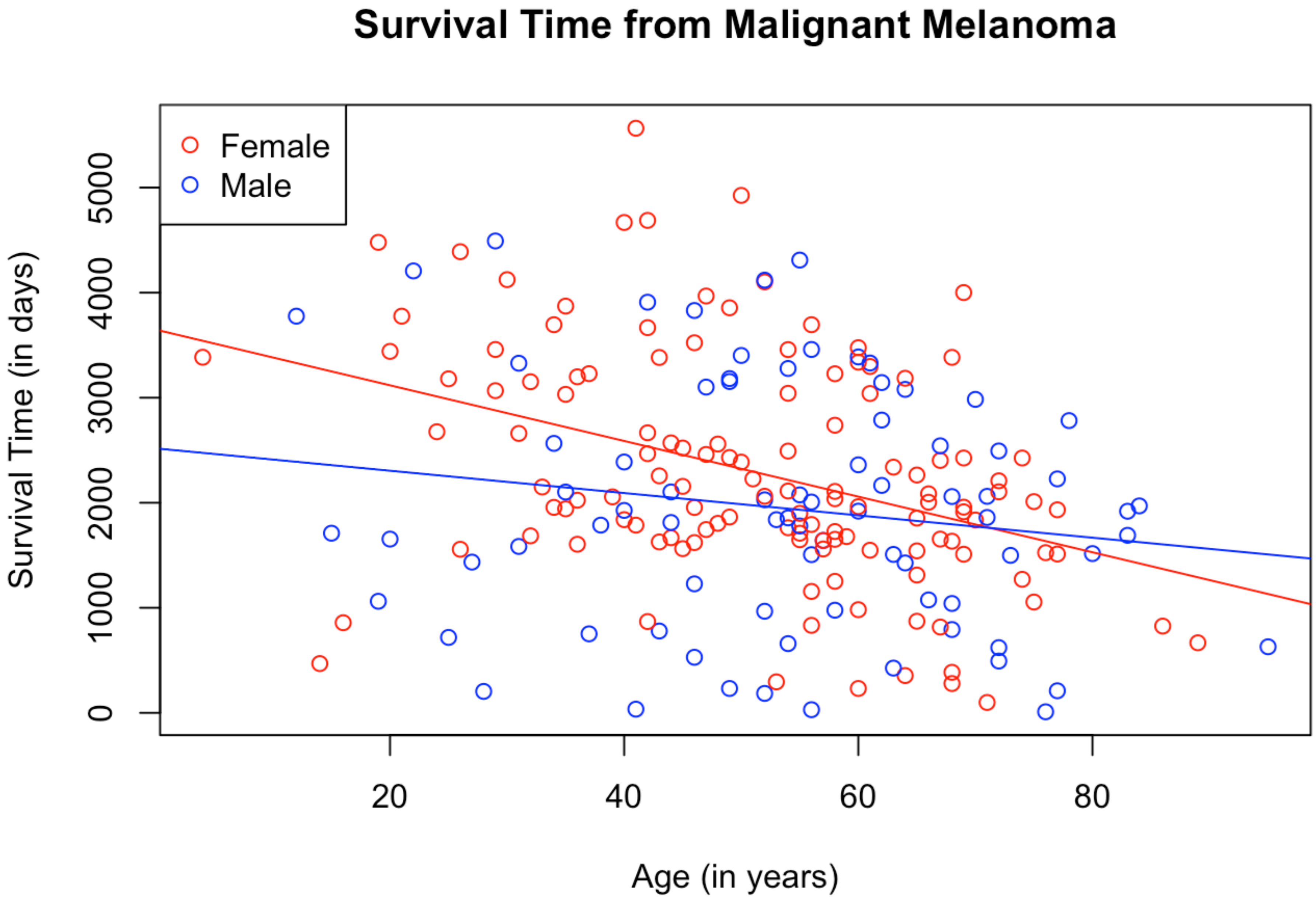
- Group the ML algorithms into groups based on their function
- There is no universal family structure, some algorithms can be placed into multiple groups
- The following examples are only a fraction of the existing algorithms

Supervised Learning

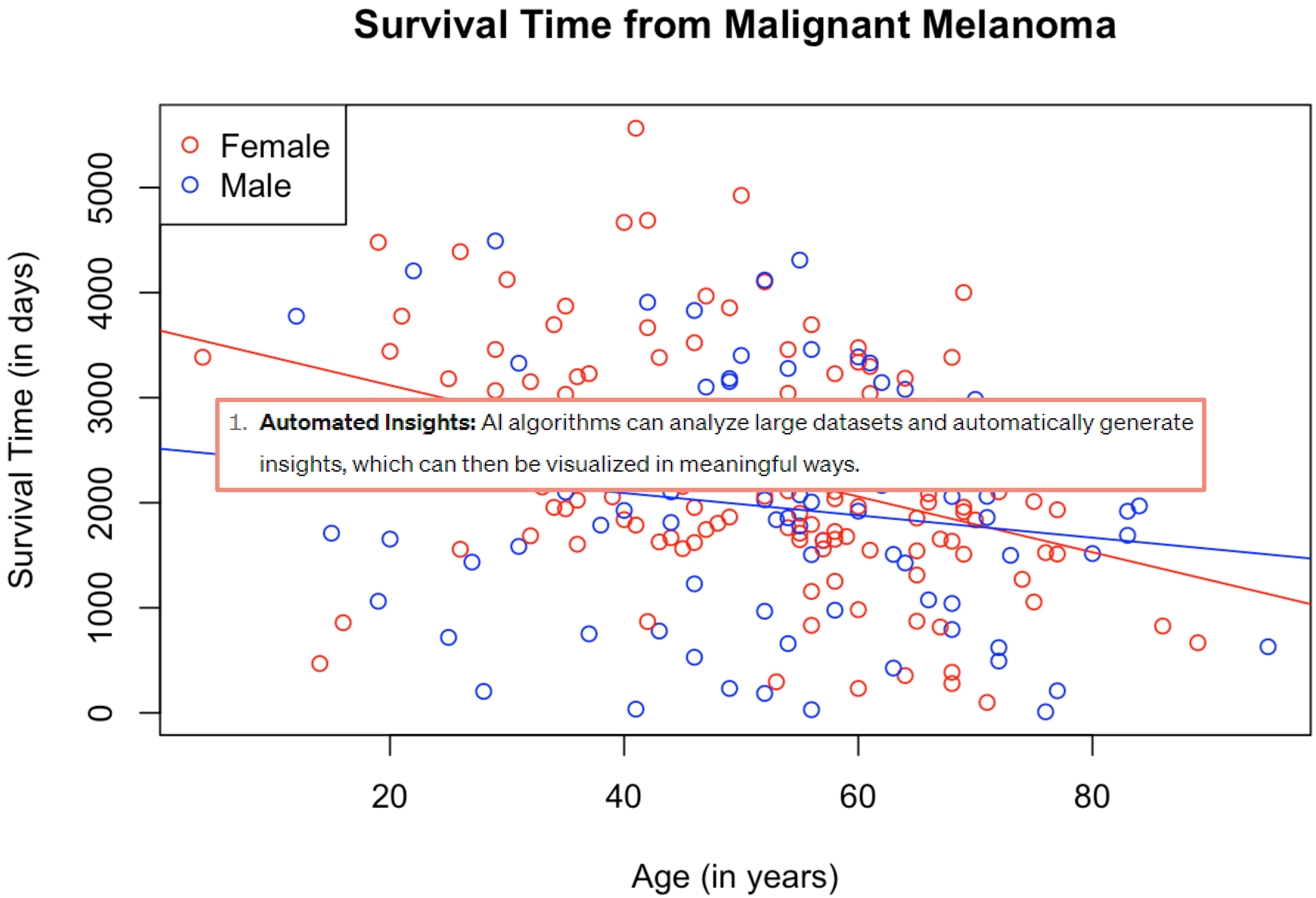




https://bookdown.org/dli/rguide/scatterplots-and-best-fit-lines-two-sets.html



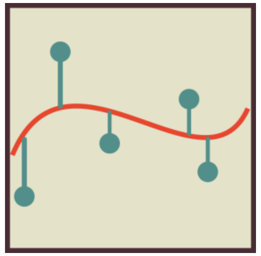
https://bookdown.org/dli/rguide/scatterplots-and-best-fit-lines-two-sets.html



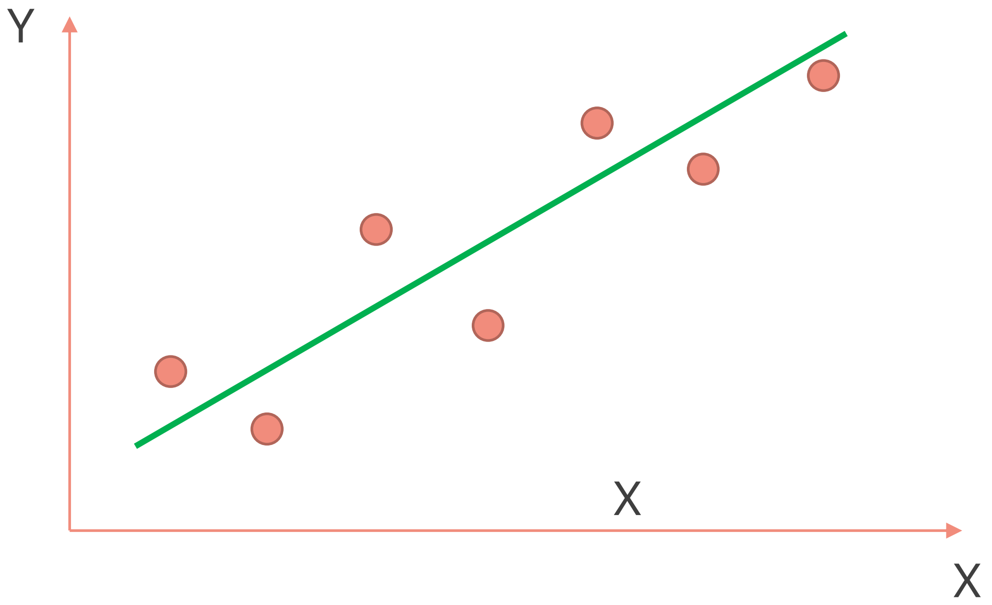
1. **Automated Insights:** AI algorithms can analyze large datasets and automatically generate insights, which can then be visualized in meaningful ways.

https://bookdown.org/dli/rguide/scatterplots-and-best-fit-lines-two-sets.html

Regression Algorithms



- Modeling the relationship between variables
  - Uses one (or multiple) independent variable
  - Tries to explain or predict the outcome of the dependent variable
  - Predict sales for a company based on weather, previous sales, GDP growth, etc.
- Iteratively refined using a measure of error in the prediction made by the model
- Examples:
  - least squares regression
  - linear regression
  - step-wise regression



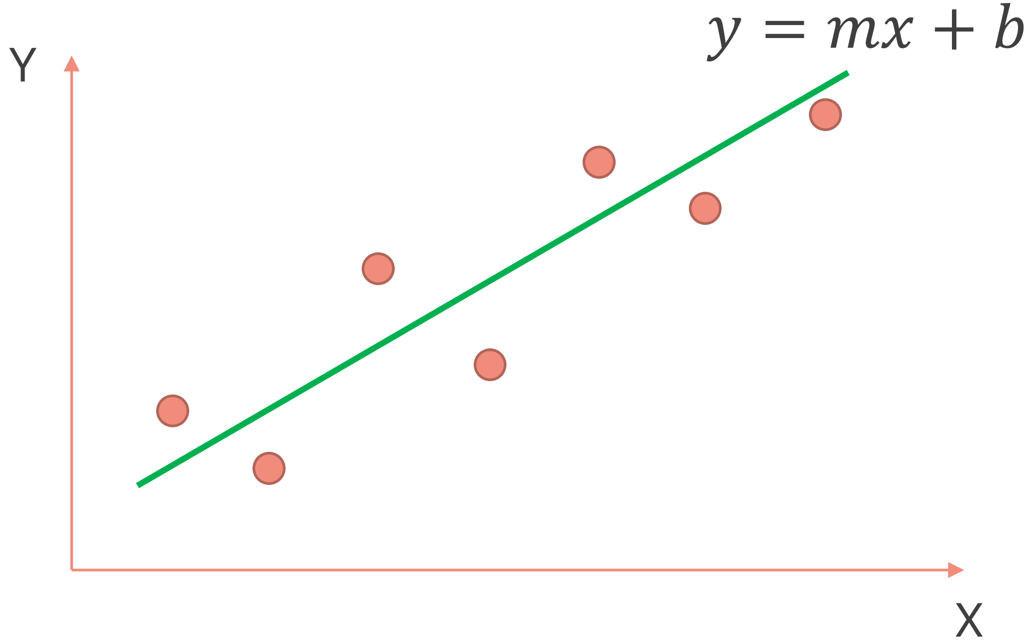
Least Squares Regression

X	1	2	3	4	5	6	7
Y	2	1	5	3	7	6	8

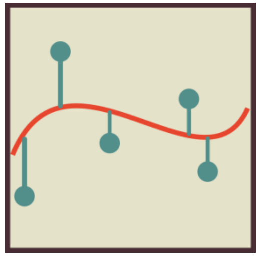
X	Y	XY	X^2
1	2	2	1
2	1	2	4
3	5	15	9
4	3	12	16
5	7	35	25
6	6	36	36
7	8	56	49
28	**32**	**158**	**140**

$$m = \frac{n * \sum xy \ - \sum x * \sum y}{n * \sum x^2 - (\sum x)^2} = \frac{7 * 158 \ - 28 * 32}{7 * 140 - 28^2} = 1.07143$$

$$b = \frac{\sum y \ - m * \sum x}{n} = \frac{32 \ - 1.07143 \ * 28}{7} = 0.28571$$



Least Squares Regression



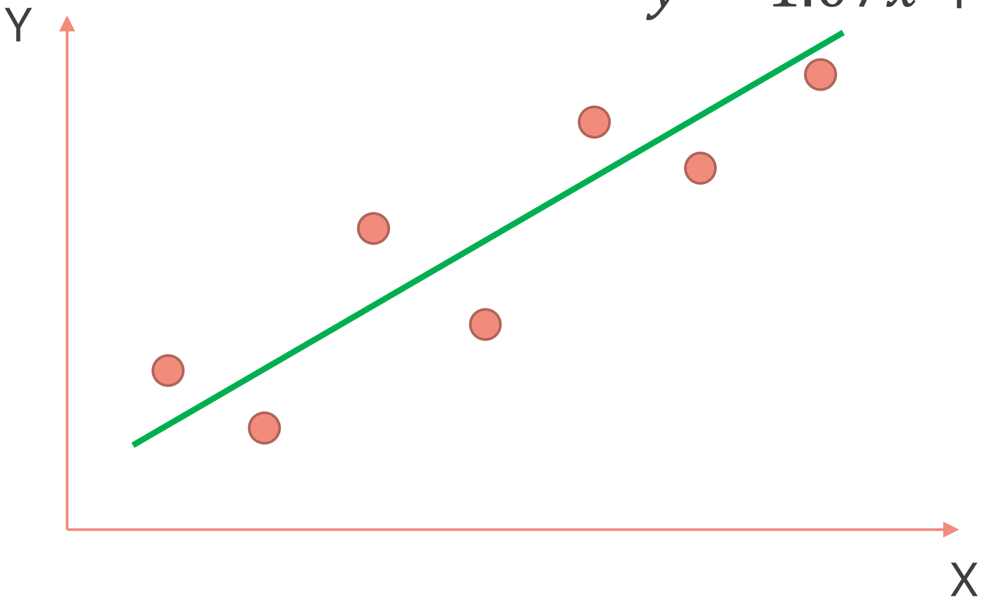
X	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Y	2	1	5	3	7	6	8

X	**Y**	**XY**	**X^2**
1	2	2	1
2	1	2	4
3	5	15	9
4	3	12	16
5	7	35	25
6	6	36	36
7	8	56	49
28	**32**	**158**	**140**

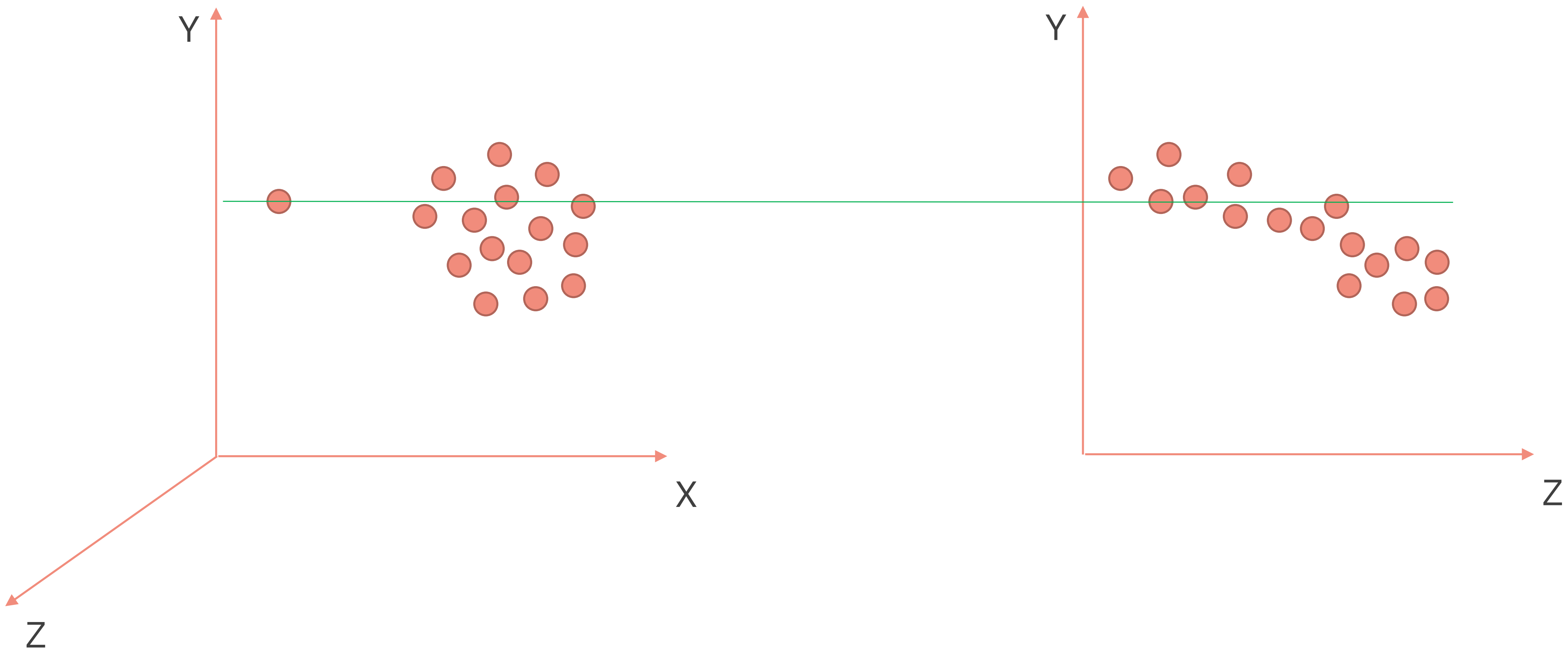
$$m = \frac{n * \sum xy - \sum x * \sum y}{n * \sum x^2 - (\sum x)^2} = \frac{7 * 158 - 28 * 32}{7 * 140 - 28^2} = 1.07143$$

$$b = \frac{\sum y - m * \sum x}{n} = \frac{32 - 1.07143 * 28}{7} = 0.28571$$

$$y = 1.07x + 0.29$$

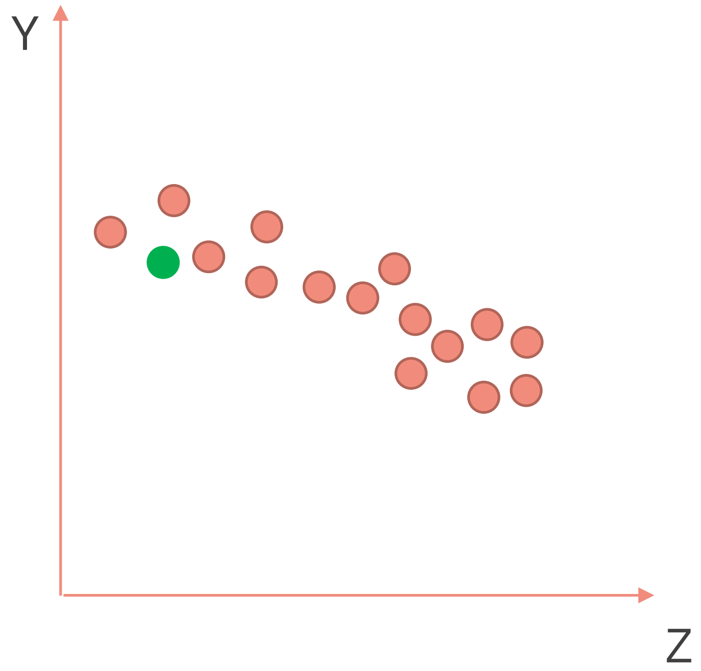


Gestalt Principles



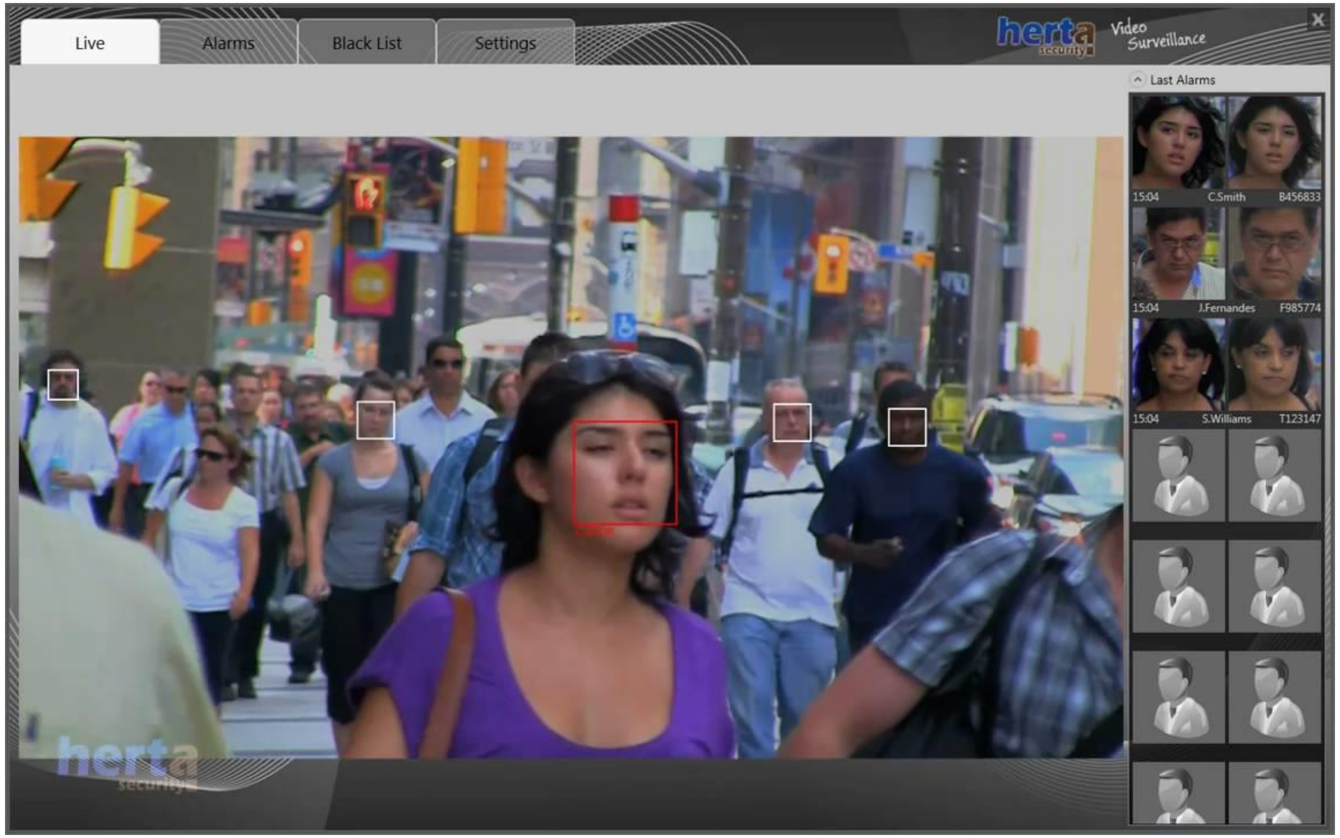
Instance-based Algorithms

- Instead of explicit generalization, compare new problems with instances seen in training
- Typically uses a database of example data
- Also called winner-take-all methods
- To reduce complexity and overfitting, instance reduction is used as preprocessing
- Examples:
  - k-nearest neighbor
  - kernel methods



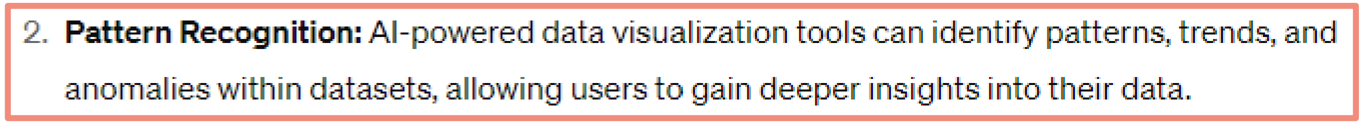
Instance-based Algorithms

- Constructs feature vectors
  - Color of eyes, distance between them, size of the nose
- Use k-NN to compare with database

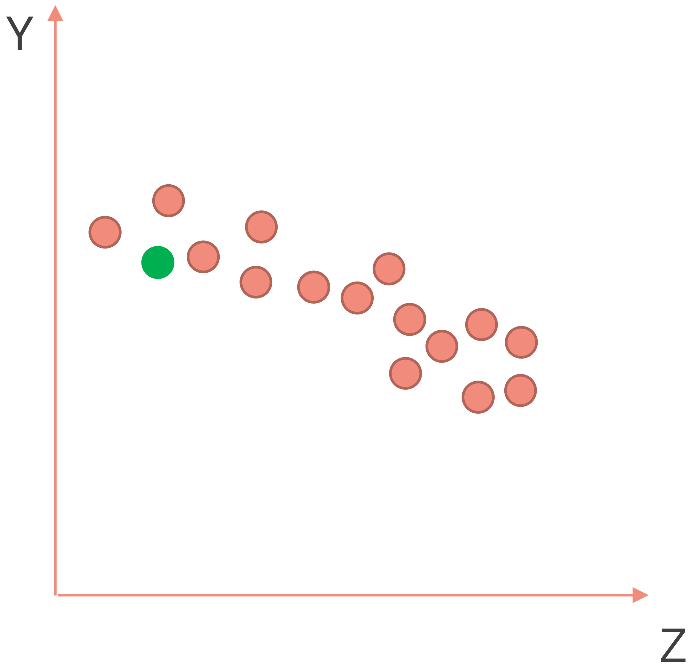


Instance-based Algorithms

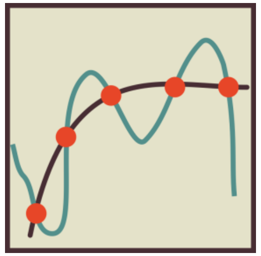
- Instead of explicit generalization, compare new problems with instances seen in training
- Typically uses a database of example data
- Also called winner-take-all methods
- To reduce complexity and overfitting, instance reduction is used as preprocessing
- Examples:
  - k-nearest neighbor
  - kernel methods



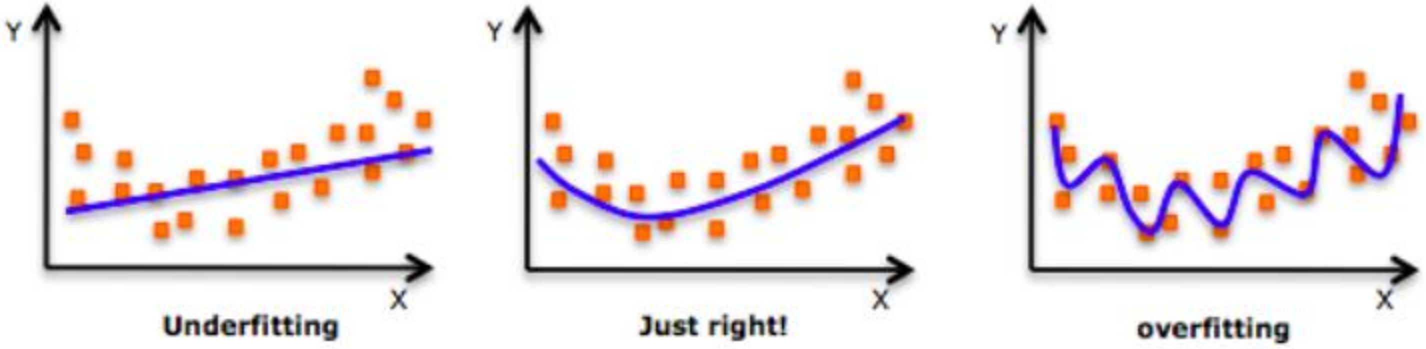
2. **Pattern Recognition:** AI-powered data visualization tools can identify patterns, trends, and anomalies within datasets, allowing users to gain deeper insights into their data.



Regularization Algorithms



- Rather an extension to other algorithms
- Introduce additional information to simplify models, reduce overfitting, create a more general algorithm
- Examples
  - Dropout regularization
  - Batch normalization
  - Early stopping



Underfitting

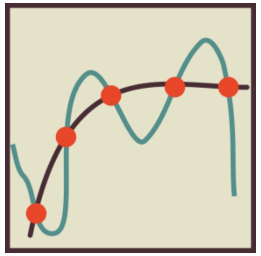
Just right!

overfitting

https://www.analyticsvidhya.com/...

Regularization Algorithms

Predict animal character

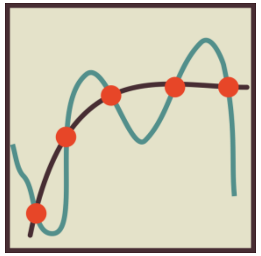


Name, color, species, size	Character
Ramses, black, cat, small	Unfriendly
Snoop, brown, dog, medium	Friendly
Boo, green, snake, small	Friendly
Lucilia, white, cat, medium	Unfriendly
Chap, yellow, dog, big	Friendly
Lis, white, dog, medium	Friendly
Napolen the third, orange, cat, small	Unfriendly
Luke, brown, snail, small	Friendly
Antonetta, black, cat, medium	Unfriendly

Rule:
Pets with names shorter than 5 letters, that are not small (except for snakes and snails) and that are not white (except for dogs) are friendly.

Regularization Algorithms

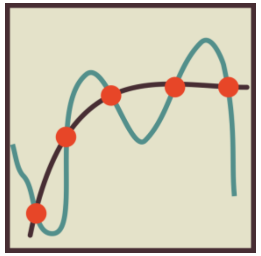
Predict animal character



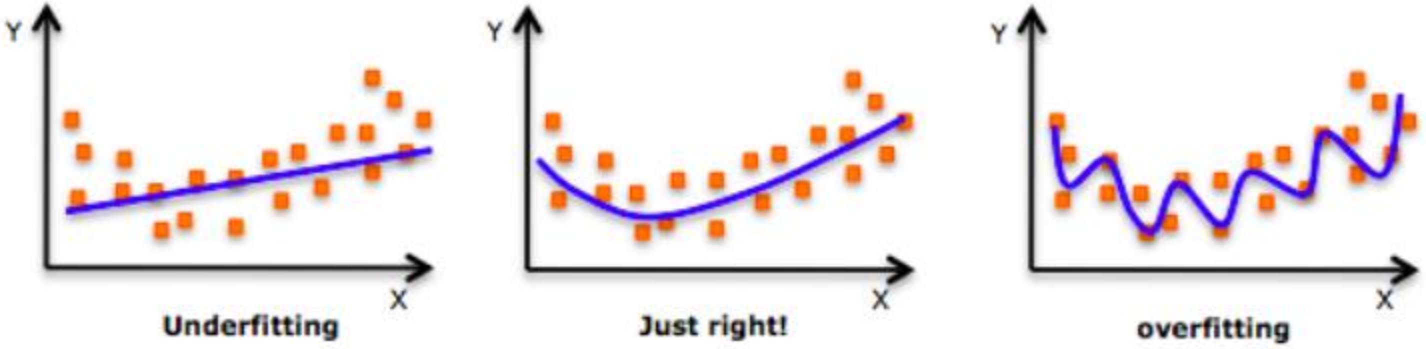
Name, color, species, size	**Character**
Ramses, black, **cat**, small	Unfriendly
Snoop, brown, **dog**, medium	Friendly
Boo, green, **snake**, small	Friendly
Lucilia, white, **cat**, medium	Unfriendly
Chap, yellow, **dog**, big	Friendly
Lis, white, **dog**, medium	Friendly
Napolen the third, orange, **cat**, small	Unfriendly
Luke, brown, **snail**, small	Friendly
Antonetta, black, **cat**, medium	Unfriendly

Rule:
Cats are unfriendly.

Regularization Algorithms

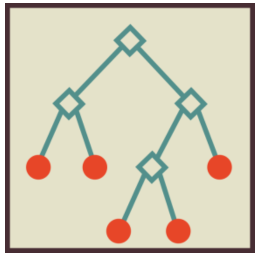


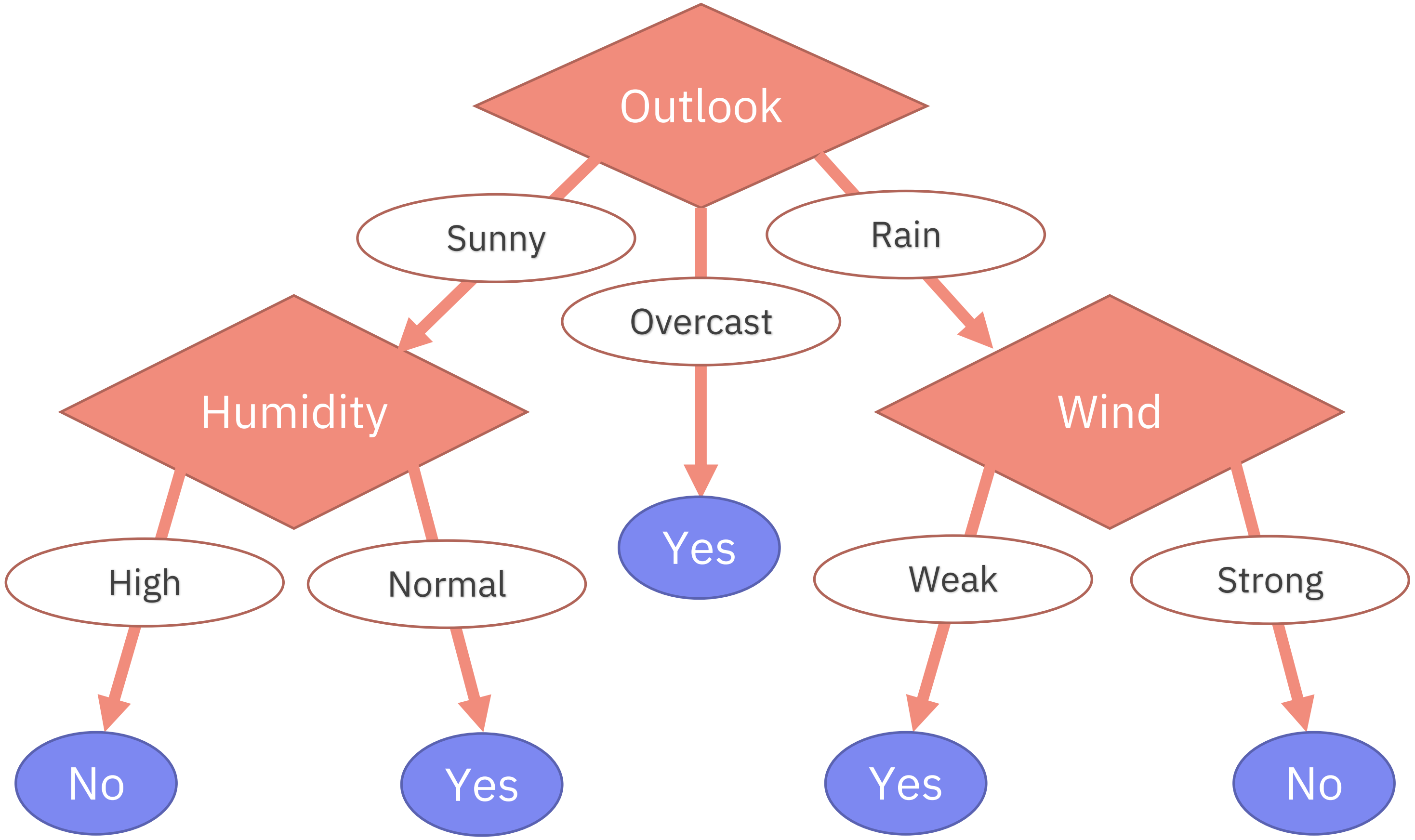
- Rather an extension to other algorithms
- Introduce additional information to simplify models, reduce overfitting, create a more general algorithm
- Examples
  - Dropout regularization
  - Batch normalization
  - Early stopping



https://www.analyticsvidhya.com/...

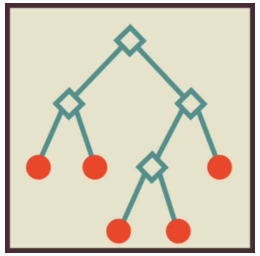
Decision Tree Algorithms





Example from (**Machine Learning, Tom Mitchell)**

Decision Tree Algorithms



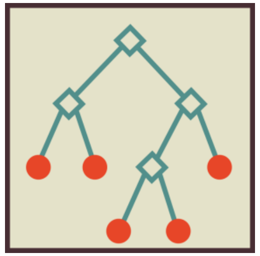
- Construct decision tree as predictive model
- Finite target variable: **classification trees**
- Continuous target variable: **regression trees**
- Requires little data preparation
- Can handle numerical and categorical data
- Examples:
  - CART (classification and regression trees)
  - Decision stump (components in ensembles)
  - Random forest (extension of bagging)

Recursive Partitioning

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Example from (**Machine Learning, Tom Mitchell)**

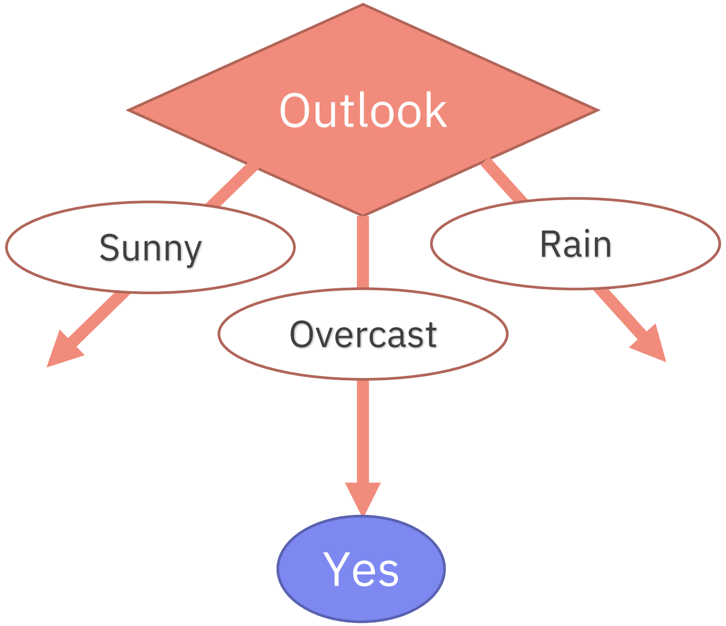
Recursive Partitioning



Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	**Overcast**	Hot	High	Weak	**Yes**
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	**Overcast**	Cool	Normal	Strong	**Yes**
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	**Overcast**	Mild	High	Strong	**Yes**
D13	**Overcast**	Hot	Normal	Weak	**Yes**
D14	Rain	Mild	High	Strong	No

Example from (**Machine Learning, Tom Mitchell**)

Recursive Partitioning



Example from (**Machine Learning, Tom Mitchell)**

Recursive Partitioning

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	**Rain**	Mild	High	Weak	**Yes**
D5	**Rain**	Cool	Normal	Weak	**Yes**
D6	**Rain**	Cool	Normal	Strong	**No**
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	**Rain**	Mild	Normal	Weak	**Yes**
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	**Rain**	Mild	High	Strong	**No**

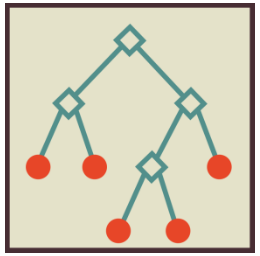
Example from (**Machine Learning, Tom Mitchell)**

Recursive Partitioning

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D4	**Rain**	Mild	High	Weak	**Yes**
D5	**Rain**	Cool	Normal	Weak	**Yes**
D6	**Rain**	Cool	Normal	Strong	**No**
D10	**Rain**	Mild	Normal	Weak	**Yes**
D14	**Rain**	Mild	High	Strong	**No**

Example from (**Machine Learning, Tom Mitchell)**

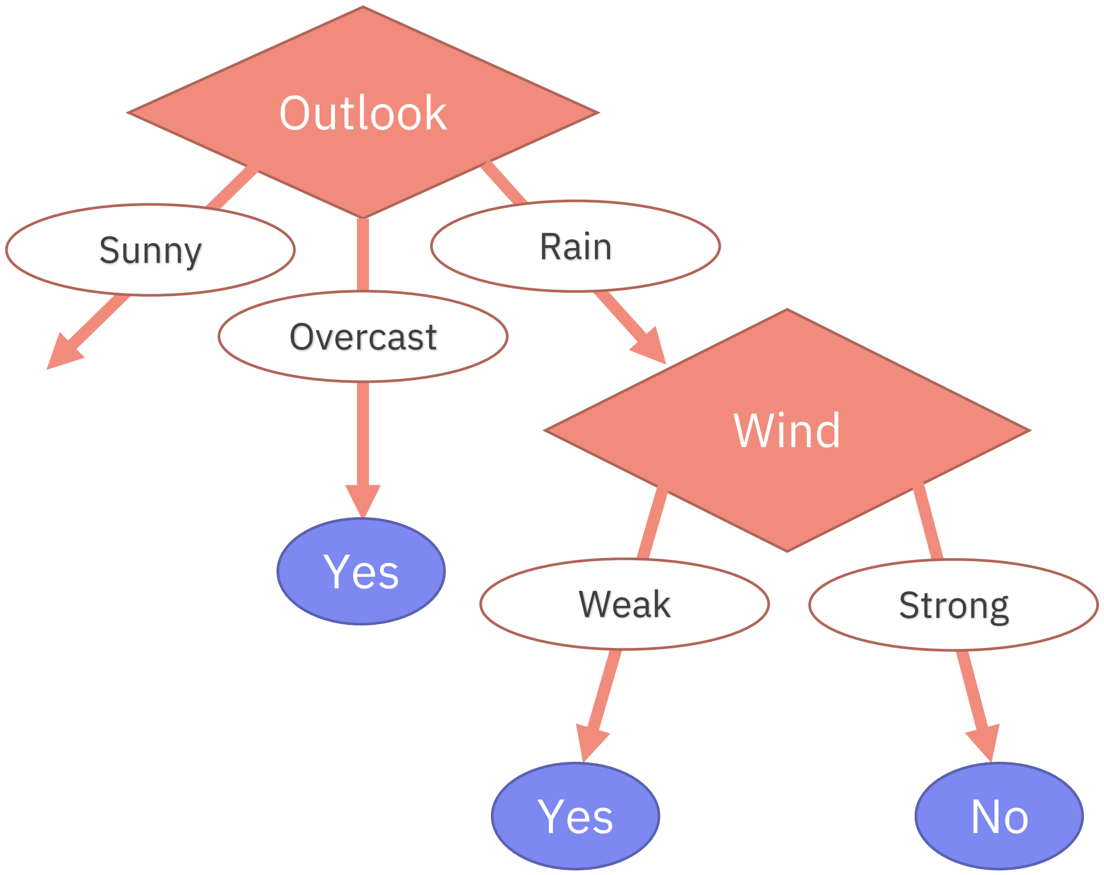
Recursive Partitioning



Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D4	**Rain**	Mild	High	**Weak**	**Yes**
D5	**Rain**	Cool	Normal	**Weak**	**Yes**
D6	**Rain**	Cool	Normal	**Strong**	**No**
D10	**Rain**	Mild	Normal	**Weak**	**Yes**
D14	**Rain**	Mild	High	**Strong**	**No**

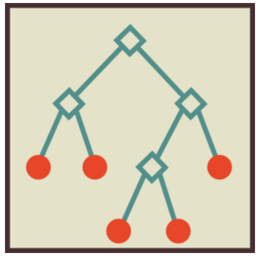
Example from (**Machine Learning, Tom Mitchell)**

Recursive Partitioning



Example from (**Machine Learning, Tom Mitchell)**

Recursive Partitioning



Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	**Sunny**	Hot	High	Weak	**No**
D2	**Sunny**	Hot	High	Strong	**No**
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	**Sunny**	Mild	High	Weak	**No**
D9	**Sunny**	Cool	Normal	Weak	**Yes**
D10	Rain	Mild	Normal	Weak	Yes
D11	**Sunny**	Mild	Normal	Strong	**Yes**
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

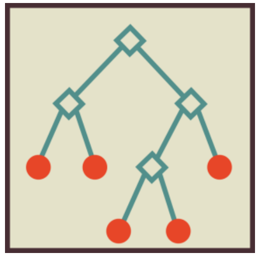
Example from (**Machine Learning, Tom Mitchell)**

Recursive Partitioning

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	**Sunny**	Hot	High	Weak	**No**
D2	**Sunny**	Hot	High	Strong	**No**
D8	**Sunny**	Mild	High	Weak	**No**
D9	**Sunny**	Cool	Normal	Weak	**Yes**
D11	**Sunny**	Mild	Normal	Strong	**Yes**

Example from (**Machine Learning, Tom Mitchell)**

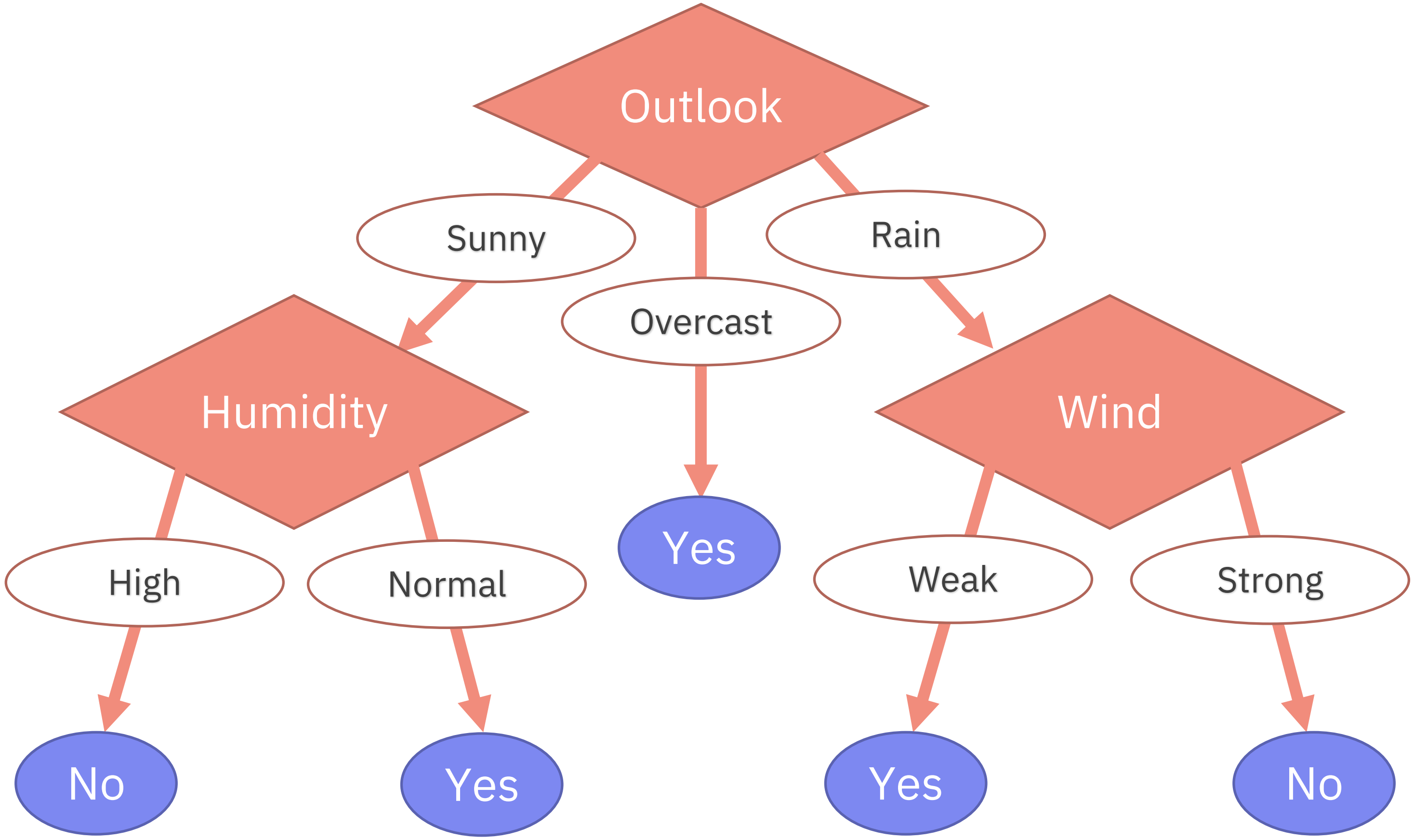
Recursive Partitioning



Day	Outlook	Temperature	Humidity	Wind	Play Tennis
D1	**Sunny**	Hot	**High**	Weak	**No**
D2	**Sunny**	Hot	**High**	Strong	**No**
D8	**Sunny**	Mild	**High**	Weak	**No**
D9	**Sunny**	Cool	**Normal**	Weak	**Yes**
D11	**Sunny**	Mild	**Normal**	Strong	**Yes**

Example from (**Machine Learning, Tom Mitchell)**

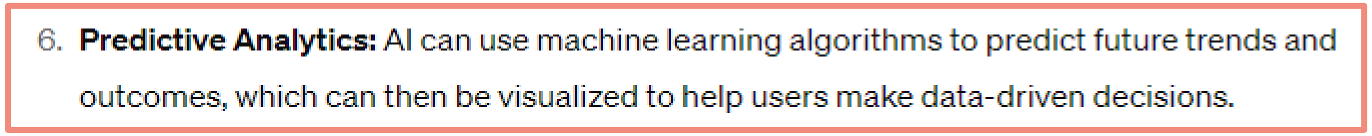
Recursive Partitioning



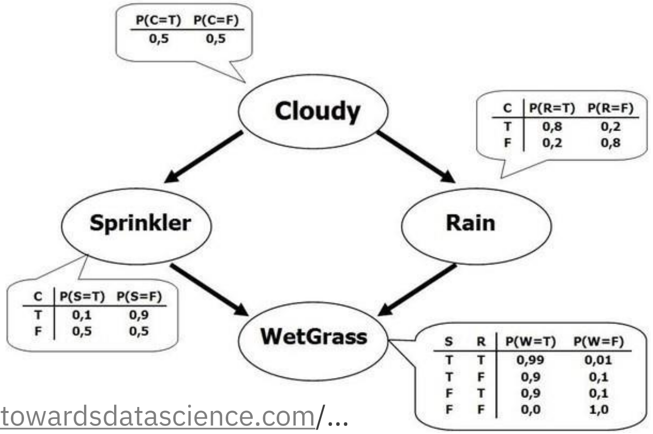
Example from (**Machine Learning, Tom Mitchell)**

Bayesian Network

- Classification based on Bayes’ theorem
- $P(\mathrm{A}|\mathrm{B}) = \frac{P(B|A)\,P(A)}{P(B)}$
- Very fast, real-time prediction
- Explainable = used in medicine
- Simplistic: presence of feature in a class is unrelated to presence of any other feature
    - A fruit is an apple if it is round, red, 7cm in diameter
    - Cancer: tobacco use, alcohol, unhealthy diet, excess body weight, physical inactivity
- Examples:
    - Gaussian Bayes (normal distribution of features)
    - Bernoulli Bayes (binary features)



6. **Predictive Analytics:** AI can use machine learning algorithms to predict future trends and outcomes, which can then be visualized to help users make data-driven decisions.



P(C=T)	P(C=F)
0,5	0,5

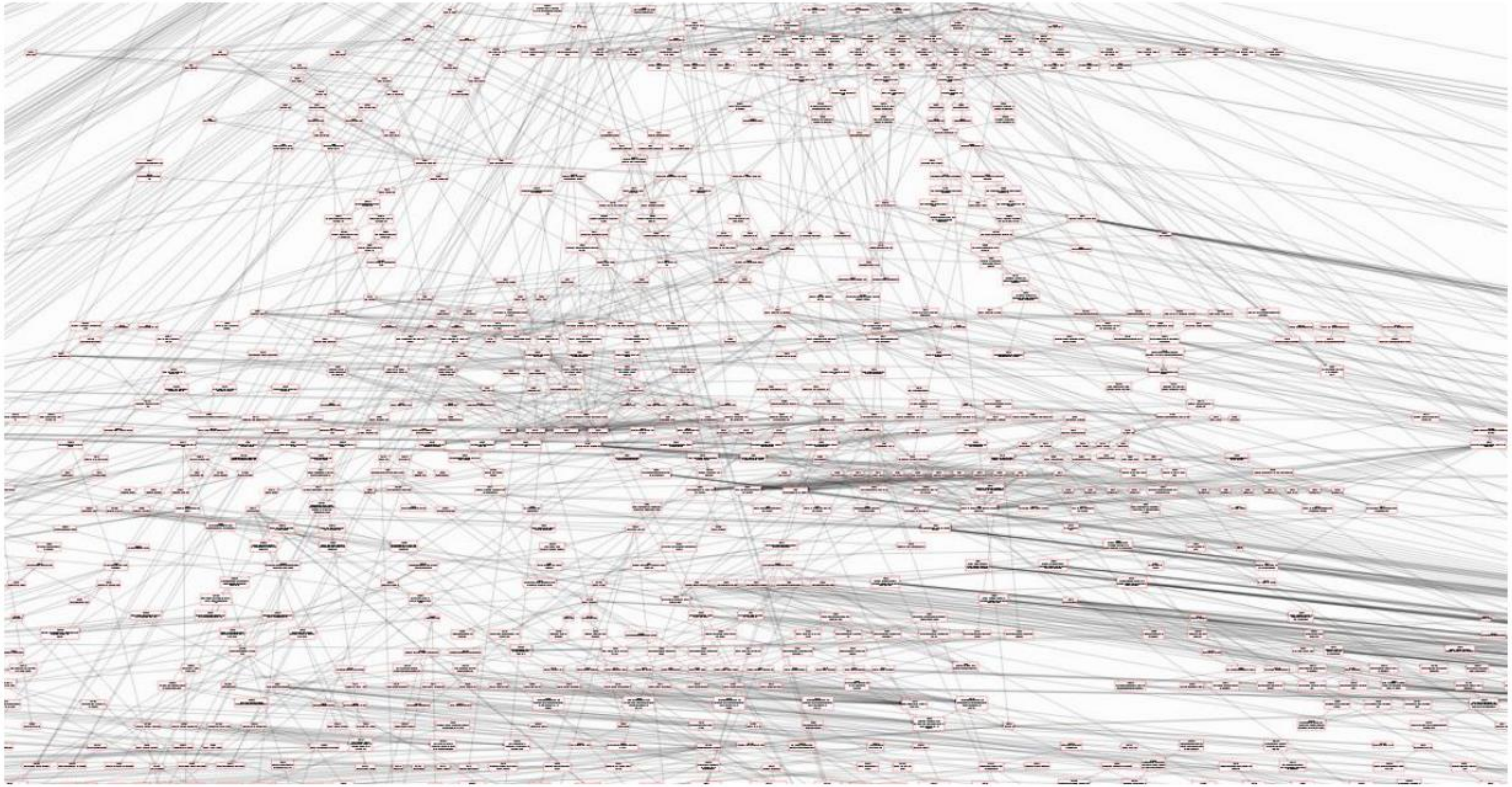
C	P(R=T)	P(R=F)
T	0,8	0,2
F	0,2	0,8

C	P(S=T)	P(S=F)
T	0,1	0,9
F	0,5	0,5

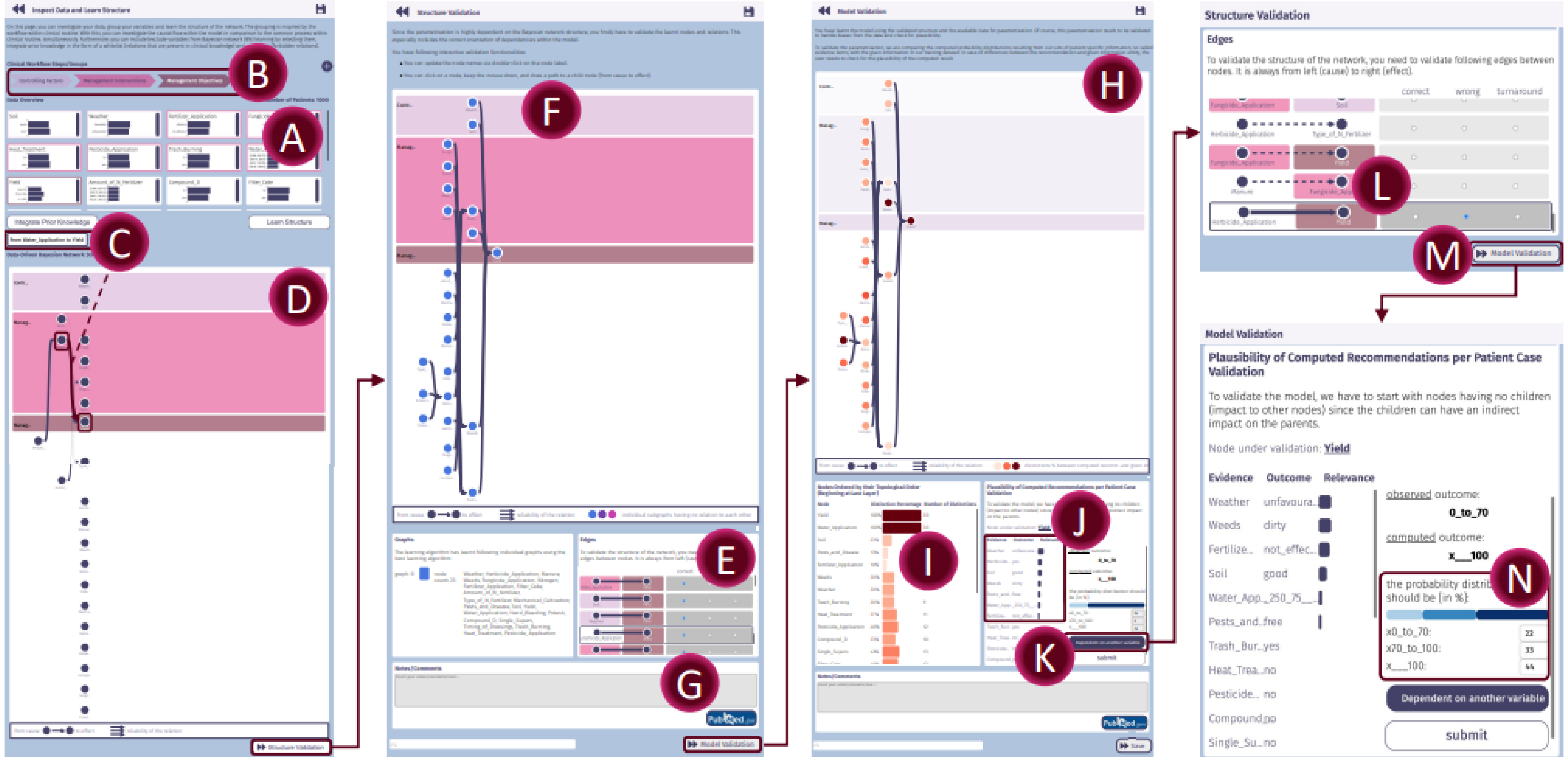
S	R	P(W=T)	P(W=F)
T	T	0,99	0,01
T	F	0,9	0,1
F	T	0,9	0,1
F	F	0,0	1,0

https://towardsdatascience.com/...

Challanges



Challanges



Structure Validation

Edges

To validate the structure of the network, you need to validate following edges between nodes. It is always from left (cause) to right (effect).

Model Validation

Plausibility of Computed Recommendations per Patient Case Validation

To validate the model, we have to start with nodes having no children (impact to other nodes) since the children can have an indirect impact on the parents.

Node under validation: **Yield**

Evidence	Outcome	Relevance
Weather	unfavoura..	
Weeds	dirty	
Fertilize..	not_effec..	
Soil	good	
Water_App..	250_75_..	
Pests_and..	free	
Trash_Bur..	yes	
Heat_Trea..	no	
Pesticide..	no	
Compound..	no	
Single_Su..	no	

observed outcome: 0_to_70

computed outcome: x__100

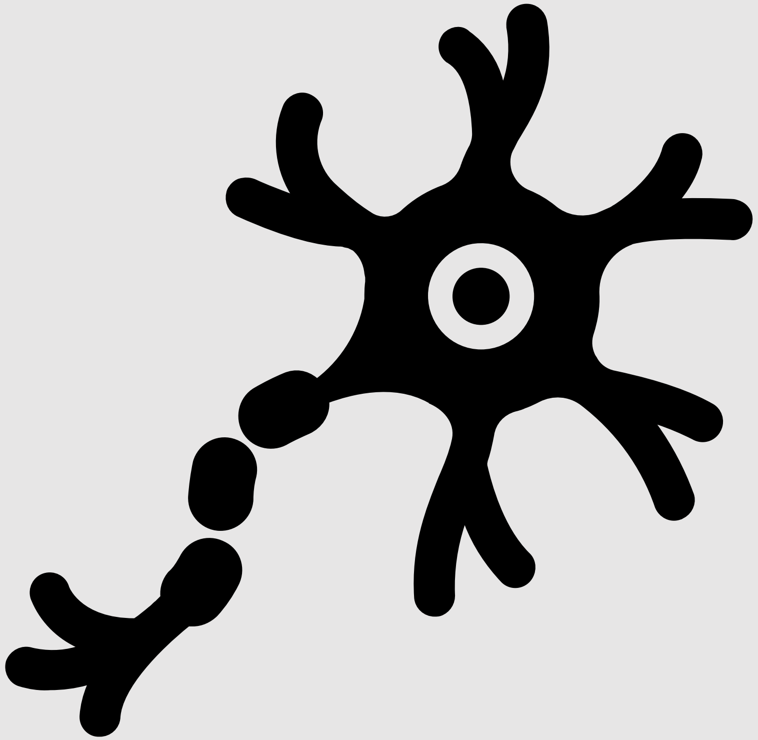
the probability distrib… should be [in %]:

x0_to_70: 22
x70_to_100: 33
x___100: 44

Visual Assistance in Development and Validation of Bayesian Networks for Clinical Decision Support

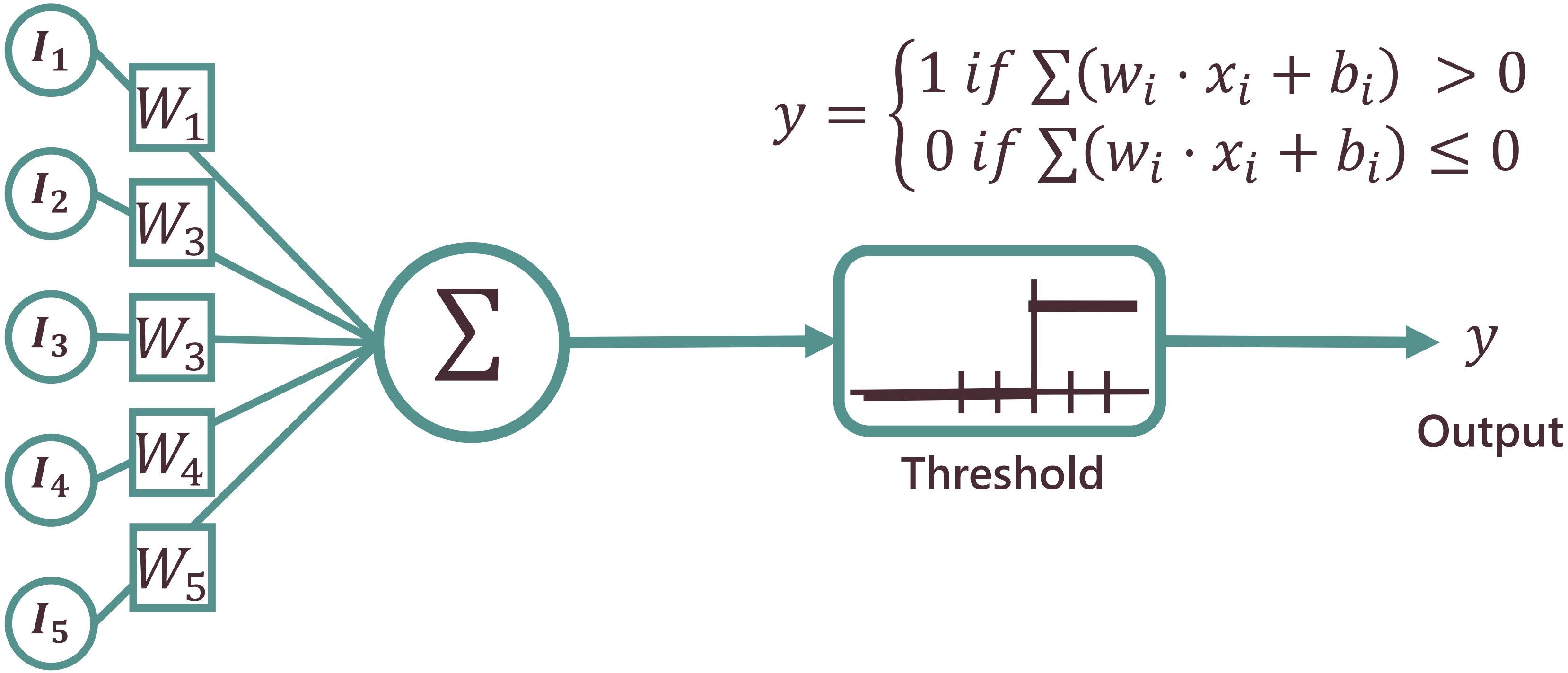
Juliane Müller-Sielaff, Seyed Behnam Beladi, Stephanie W. Vrede, Monique Meuschke, Peter J.F. Lucas, Johanna M.A. Pijnenborg, Steffen Oeltze-Jafra

Neural Networks

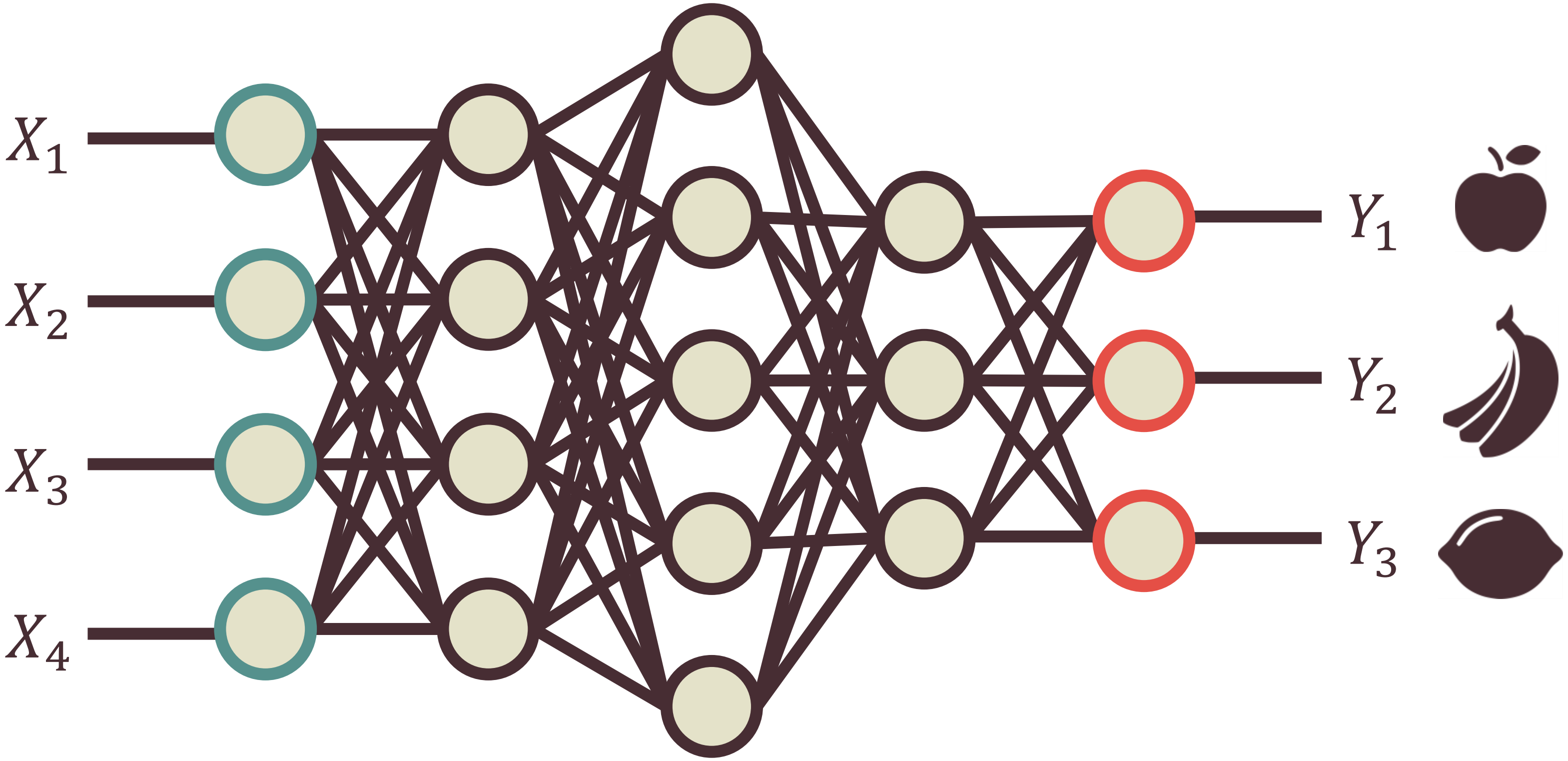


Perceptron

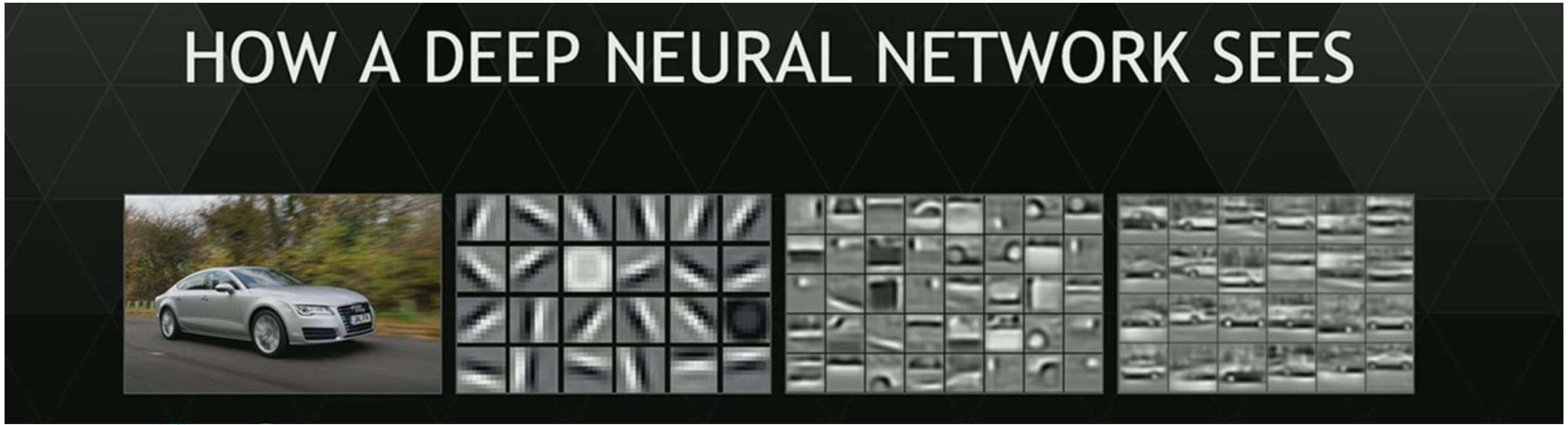
$$y = \begin{cases} 1 \; if \; \sum(w_i \cdot x_i + b_i) \; > 0 \\ 0 \; if \; \sum(w_i \cdot x_i + b_i) \leq 0 \end{cases}$$

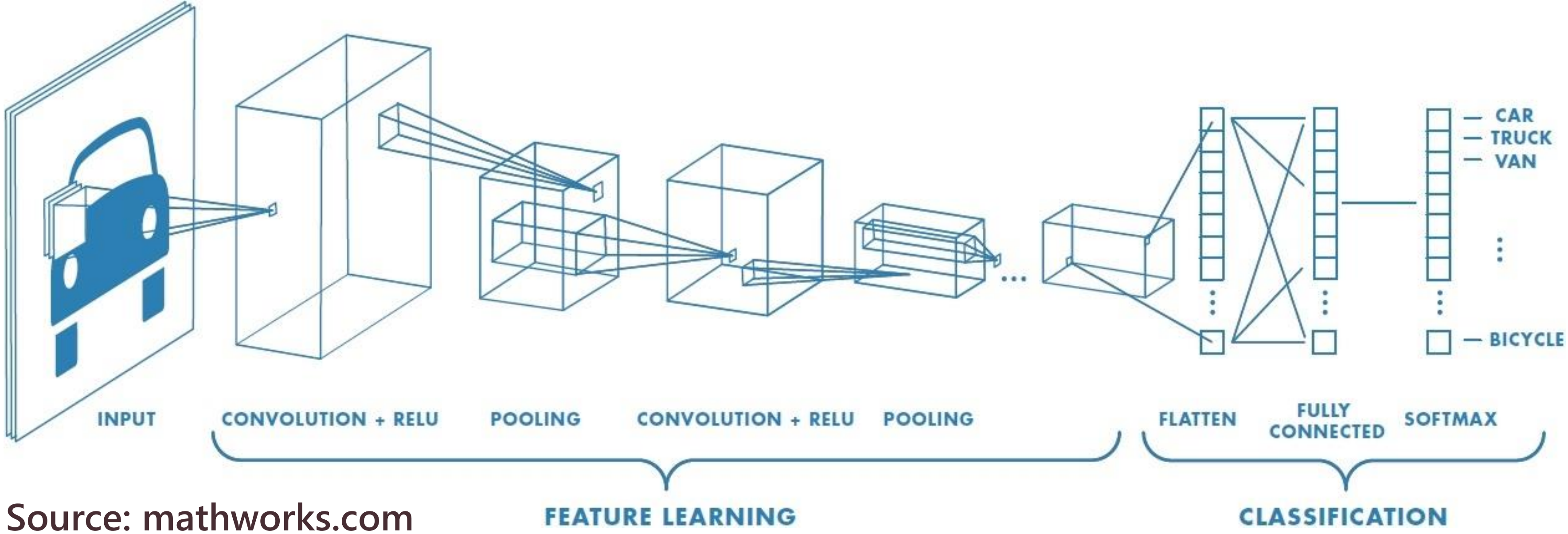


Multilayer Perceptron



Deep Learning Algorithms

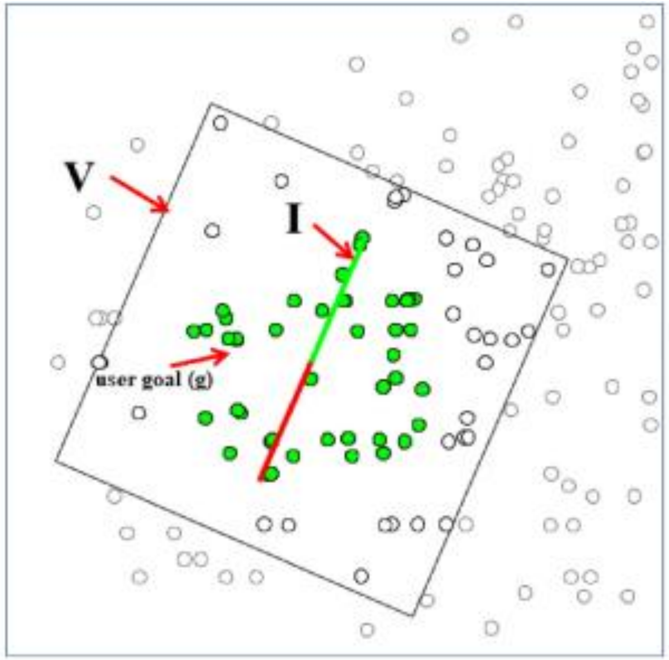


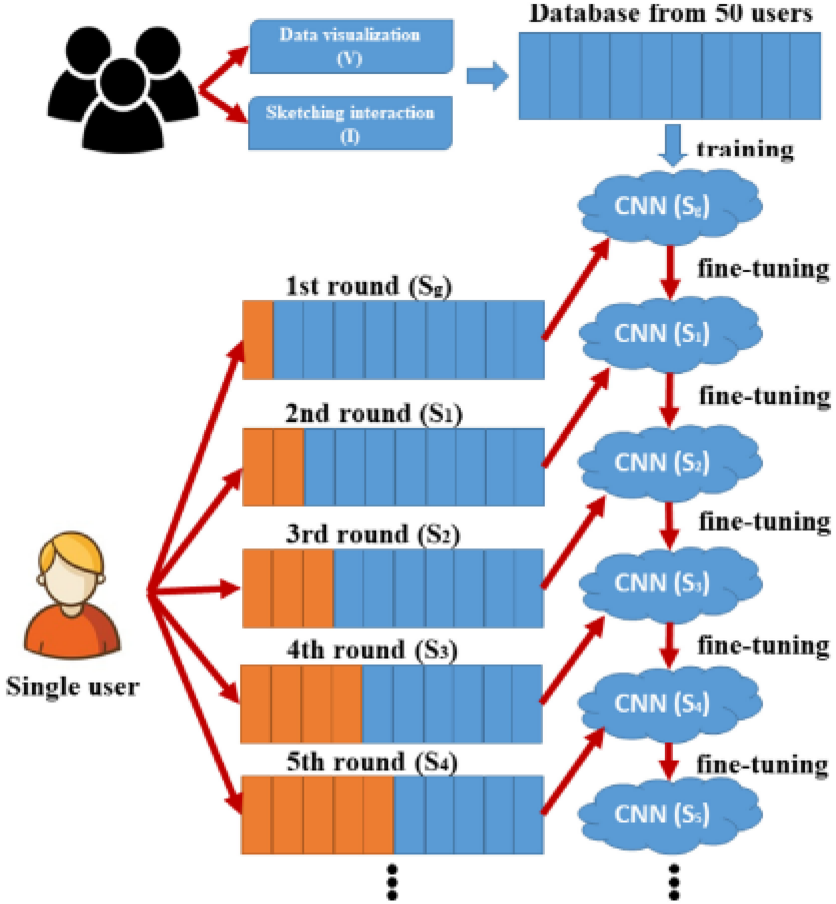


Source: mathworks.com

Personalized Sketch-Based Brushing in Scatterplots

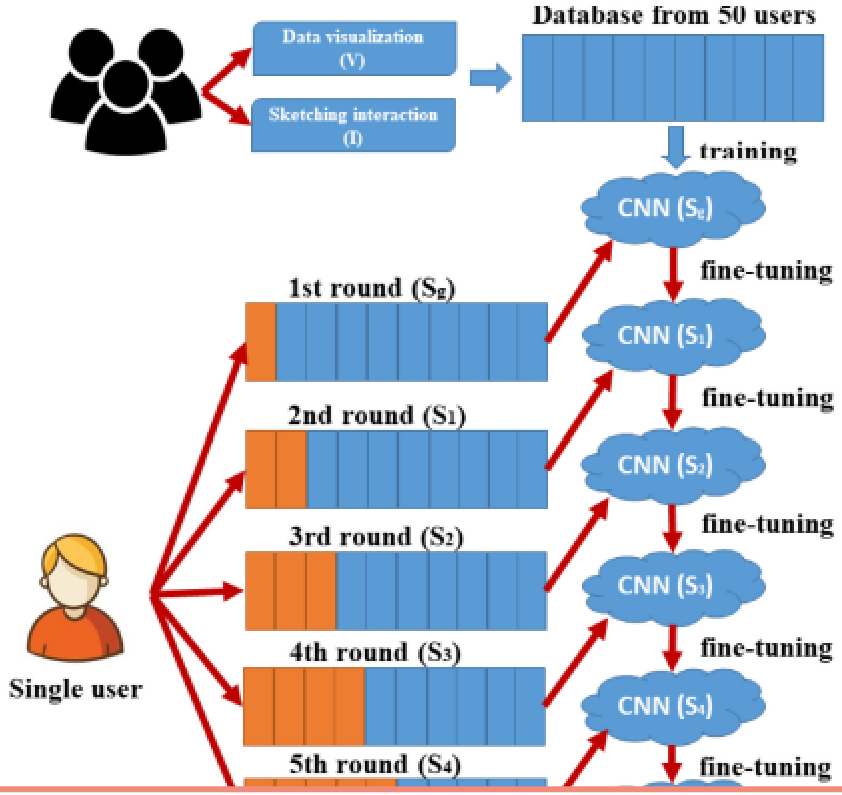
- Predicting the user’s brushing goal
  - Average brushing preference
- Improve the brushing technique while using it

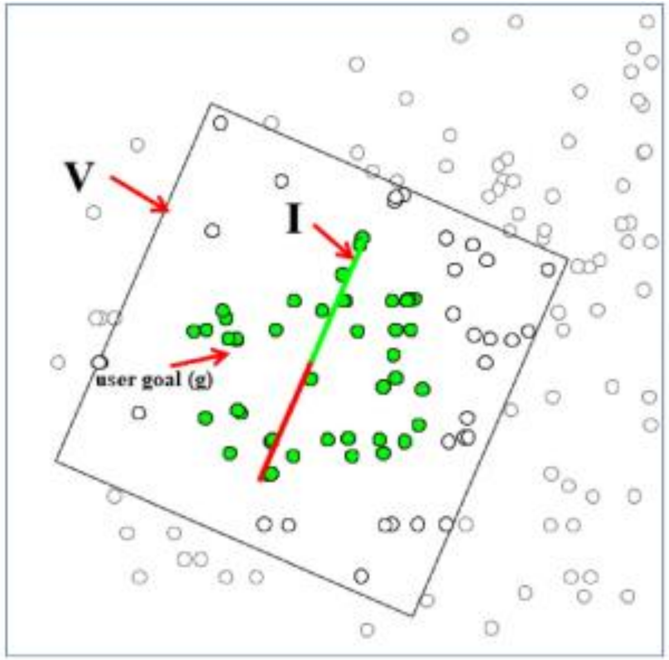


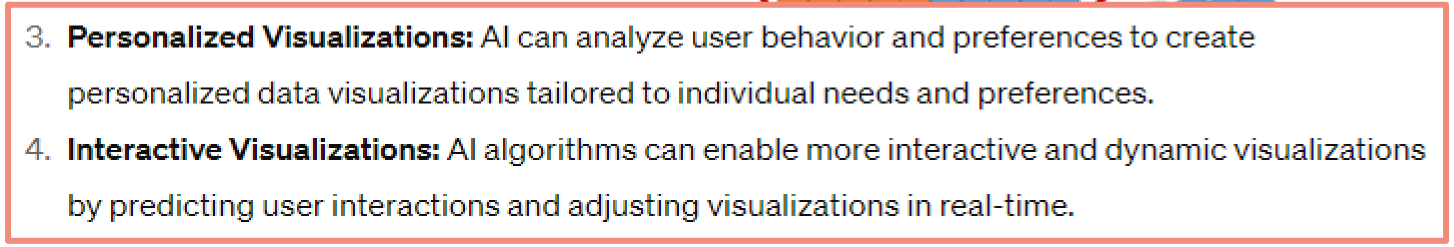


Personalized Sketch-Based Brushing in Scatterplots

- Predicting the user’s brushing goal
  - Average brushing preference
- Improve the brushing technique while using it

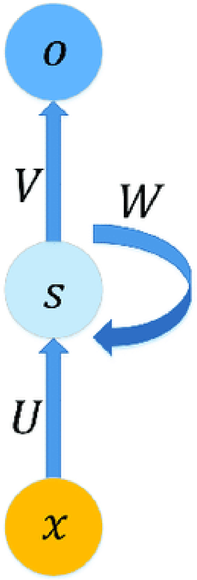





3. **Personalized Visualizations:** AI can analyze user behavior and preferences to create personalized data visualizations tailored to individual needs and preferences.
4. **Interactive Visualizations:** AI algorithms can enable more interactive and dynamic visualizations by predicting user interactions and adjusting visualizations in real-time.

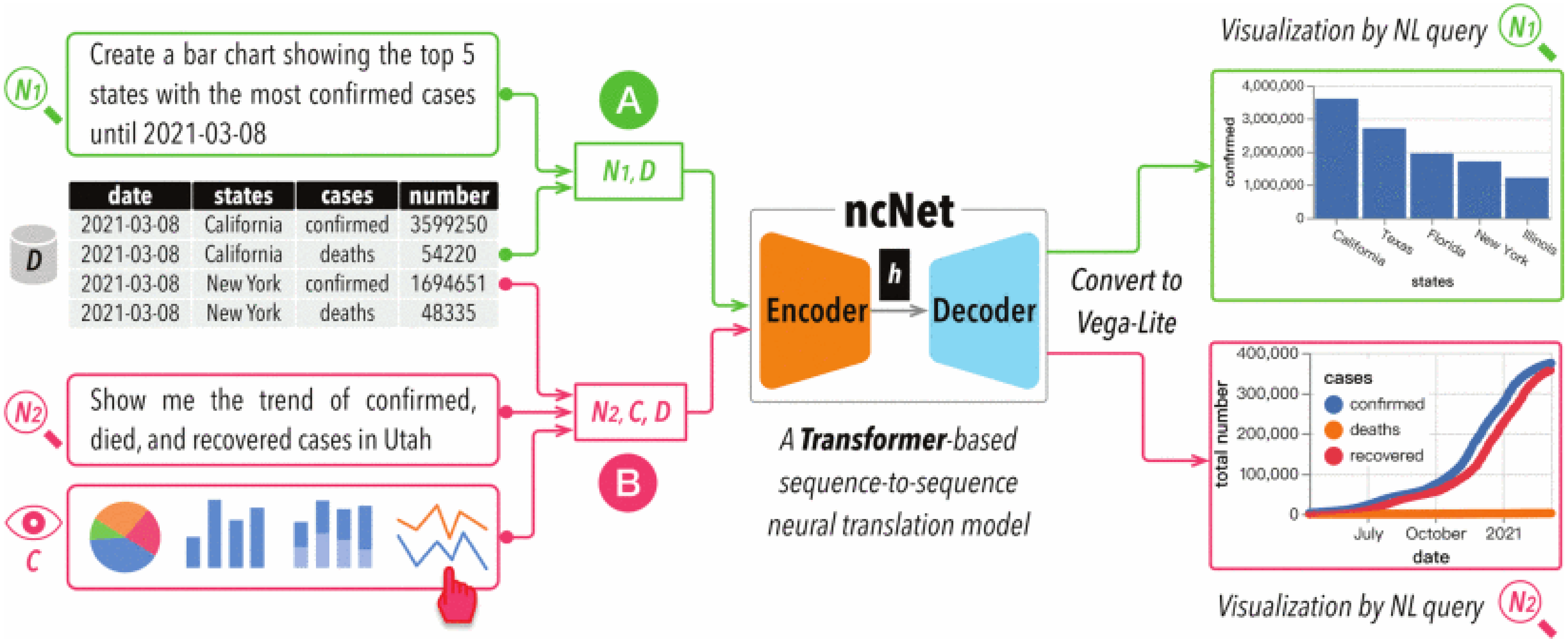
Recurrent Neural Network

- Internal memory to include the result from previous classification
- Used in cases when temporal domain is important
- Used for speech recognition



Bao, Wei, Jun Yue, and Yulei Rao. "A deep learning framework for financial time series using stacked autoencoders and long-short term memory."

Transformers



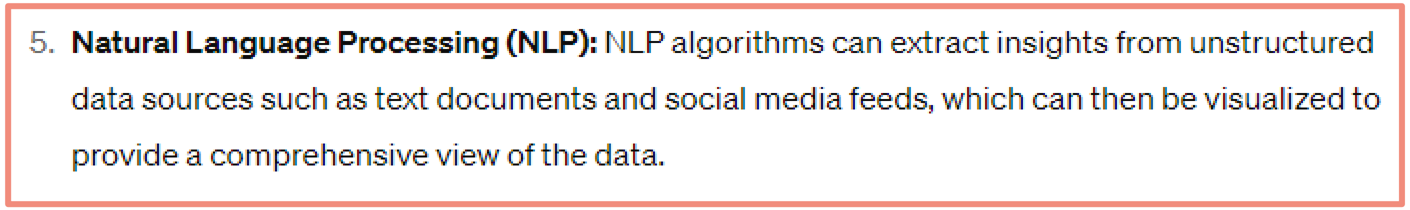
Create a bar chart showing the top 5 states with the most confirmed cases until 2021-03-08

date	states	cases	number
2021-03-08	California	confirmed	3599250
2021-03-08	California	deaths	54220
2021-03-08	New York	confirmed	1694651
2021-03-08	New York	deaths	48335

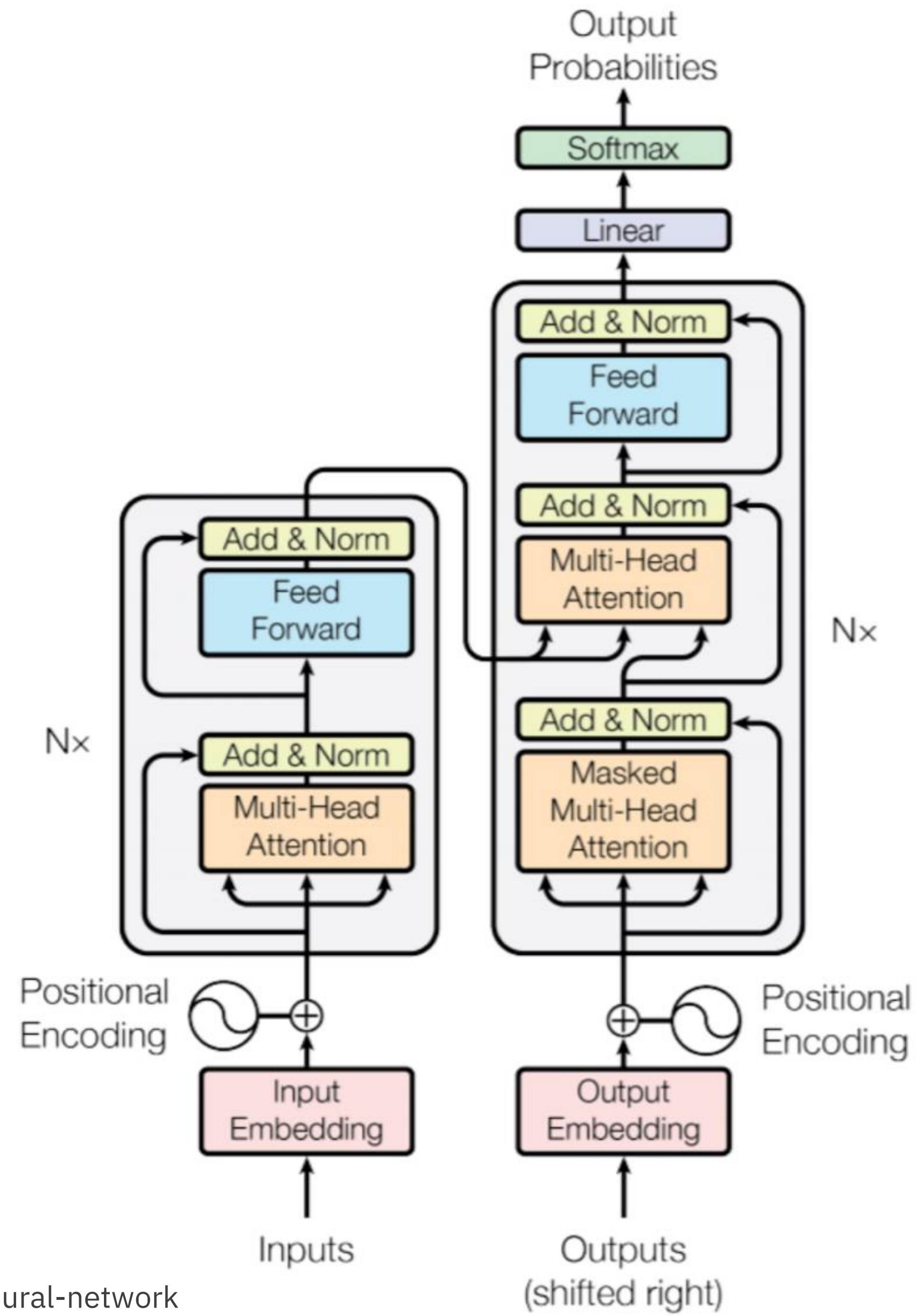
Show me the trend of confirmed, died, and recovered cases in Utah

A *Transformer*-based sequence-to-sequence neural translation model

Transformers



5. **Natural Language Processing (NLP):** NLP algorithms can extract insights from unstructured data sources such as text documents and social media feeds, which can then be visualized to provide a comprehensive view of the data.

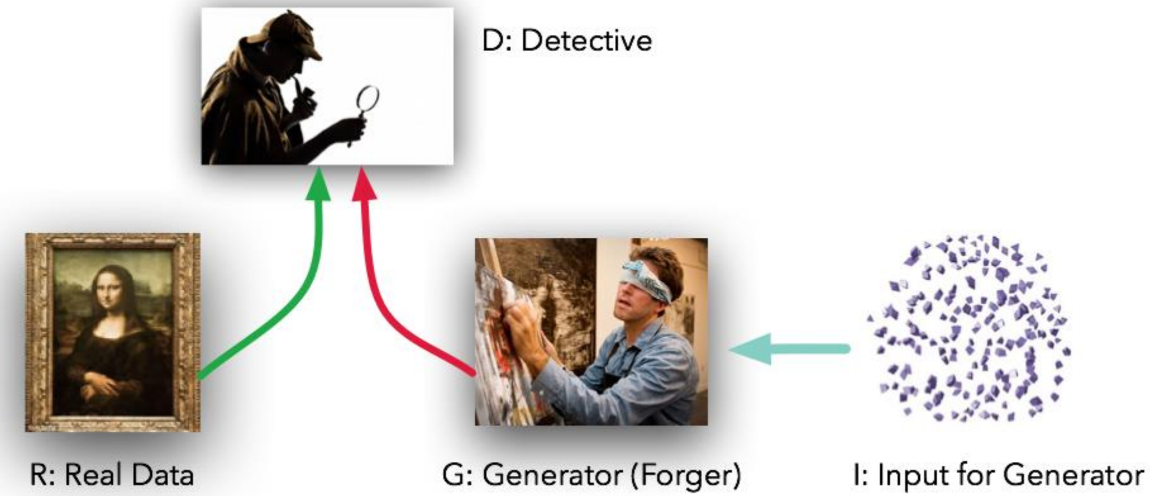


https://builtin.com/artificial-intelligence/transformer-neural-network

Generative Networks

Ian j. Goodfellow et al.

Noogle
Ian Goodfellow
• Job Title: Software Eng Intern
• Degree and University: PhD Université de Montréal
• Fun Fact About You: Took surf PE in high school
Noogler



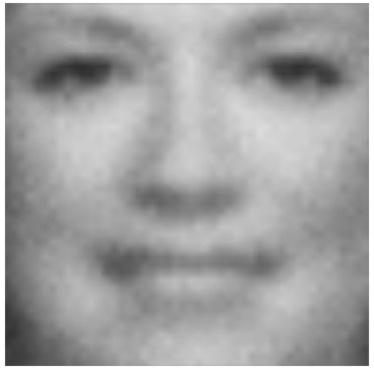
D: Detective

R: Real Data

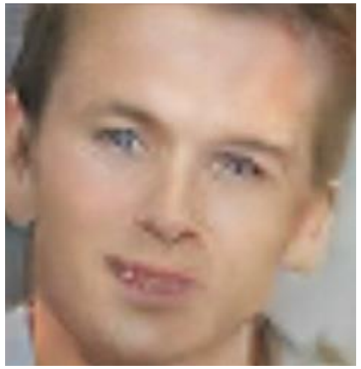
G: Generator (Forger)

I: Input for Generator

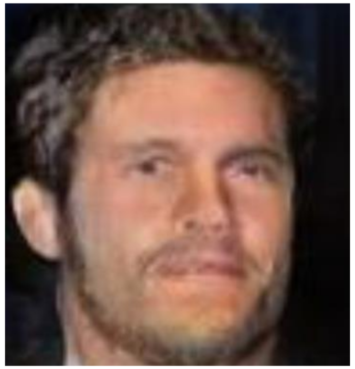
Result Progression



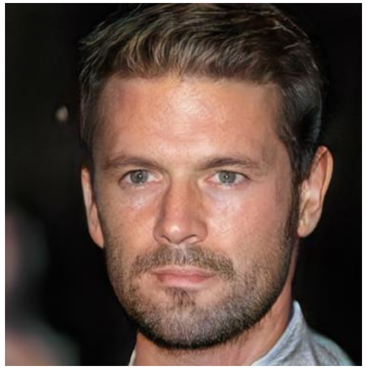
2014
Goodfellow et al.



2015
Radfort et al.



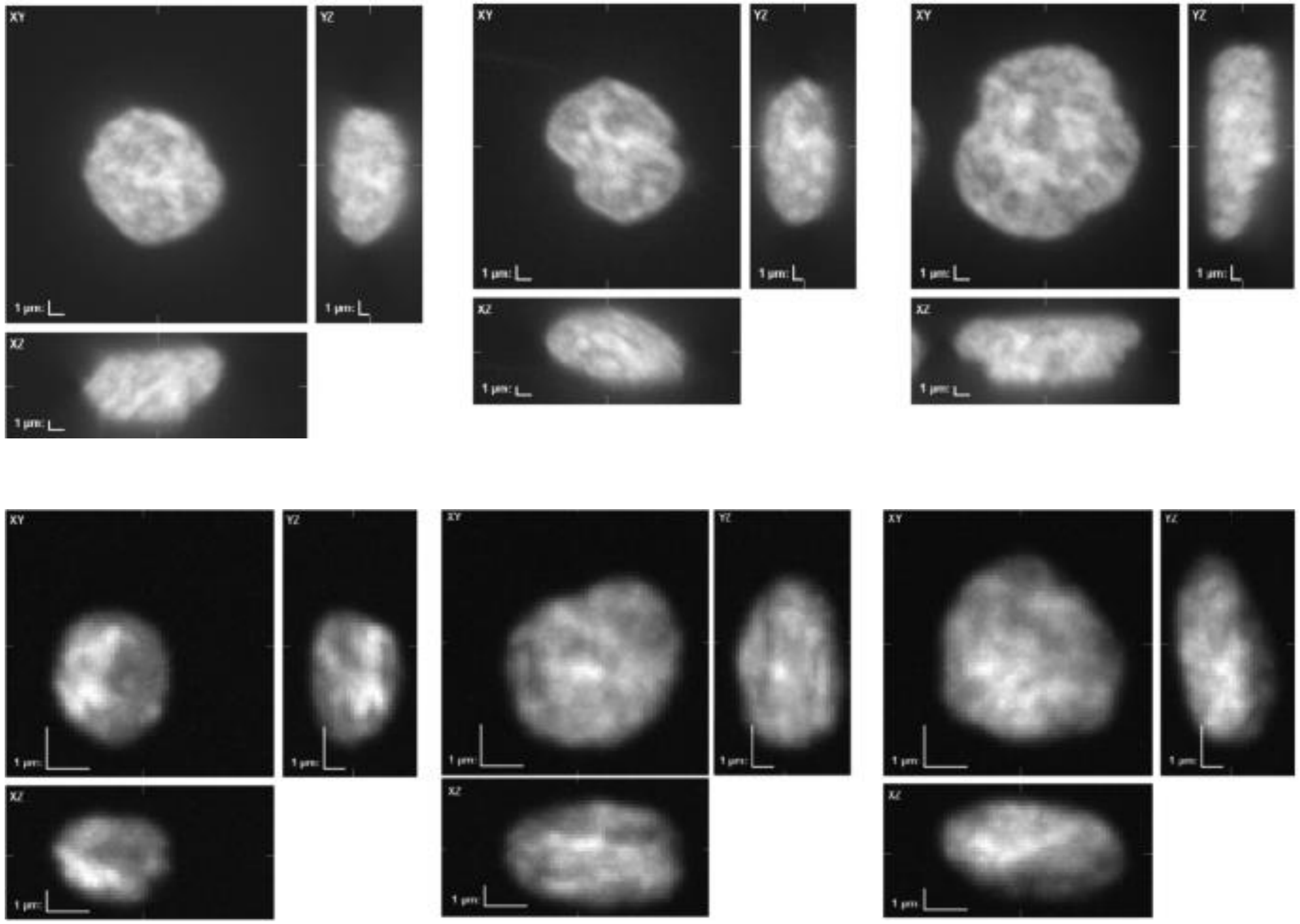
2016
Liu and Tuzel



2017
Karras et al.

Generative Modeling of Cell Shape Using 3D GANs

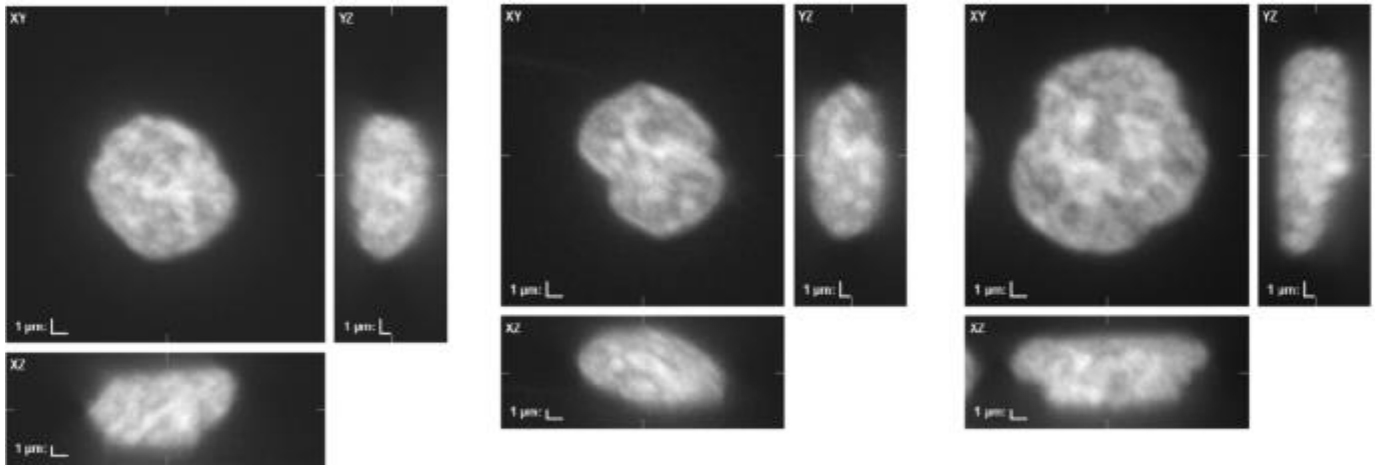
- Obtaining real data may be expensive
- Generating synthetic cellular specimens to produce suitable testing datasets

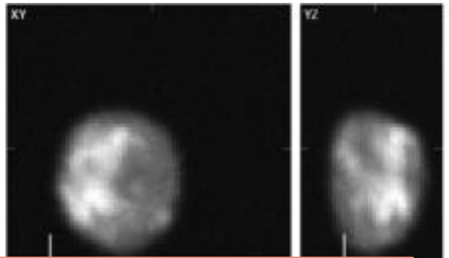


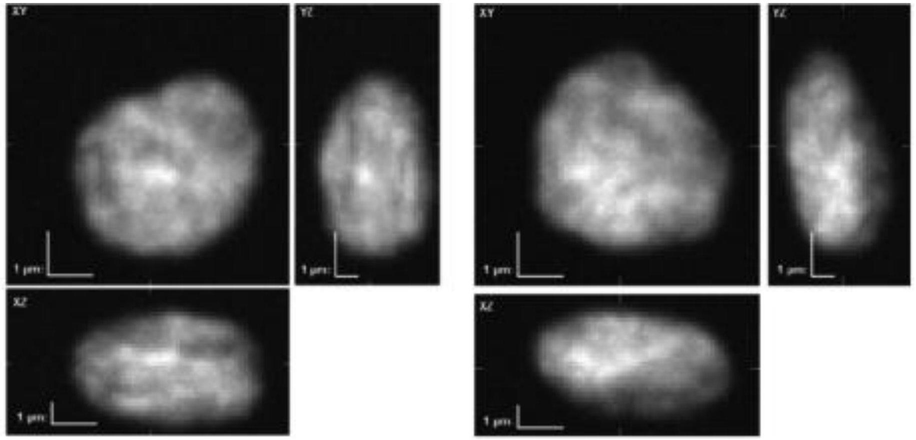
Wiesner, D., Nečasová, T., & Svoboda, D. (2019)

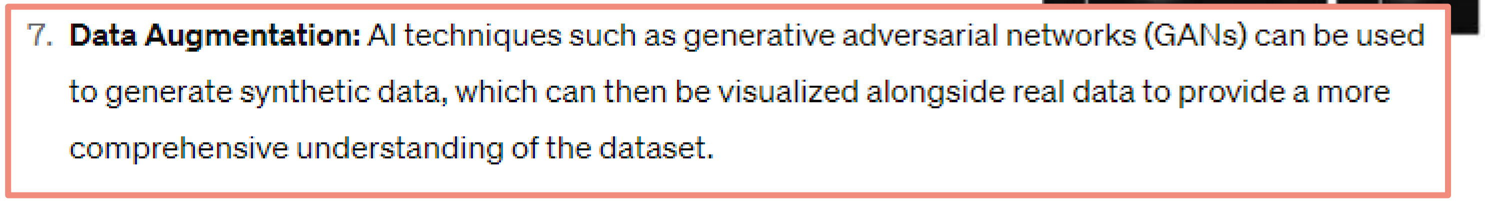
Generative Modeling of Cell Shape Using 3D GANs

- Obtaining real data may be expensive
- Generating synthetic cellular specimens to produce suitable testing datasets





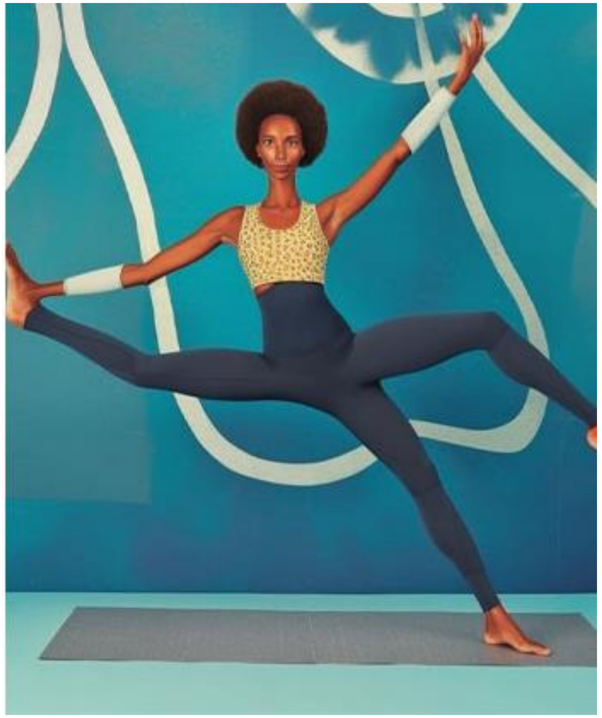



7. **Data Augmentation:** AI techniques such as generative adversarial networks (GANs) can be used to generate synthetic data, which can then be visualized alongside real data to provide a more comprehensive understanding of the dataset.

Wiesner, D., Nečasová, T., & Svoboda, D. (2019)

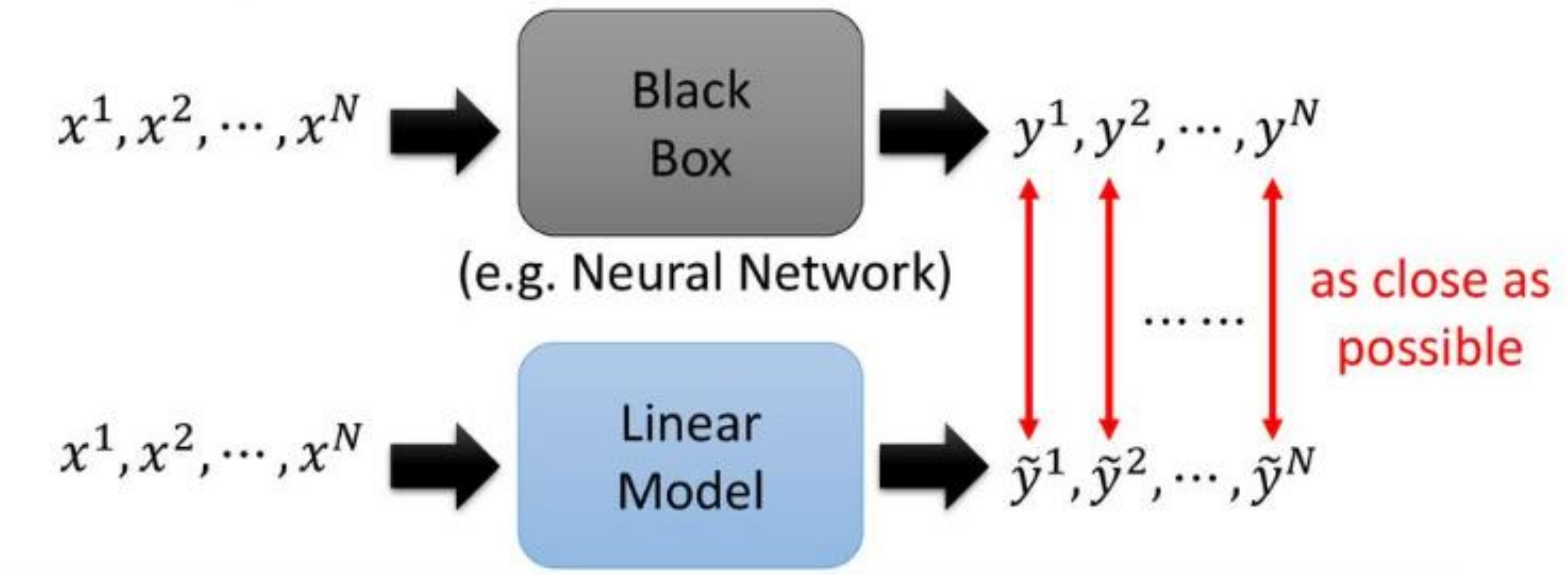
Result Progression

https://www.boredpanda.com/ai-fails/

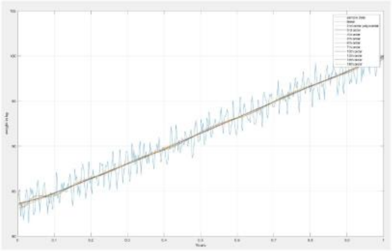


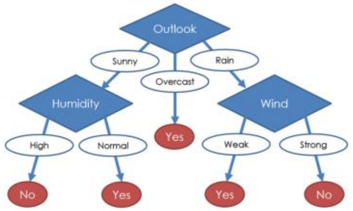
Explainable AI

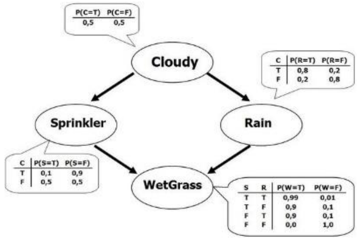
Using a Model to Explain Another



Source: Hung-yi Lee

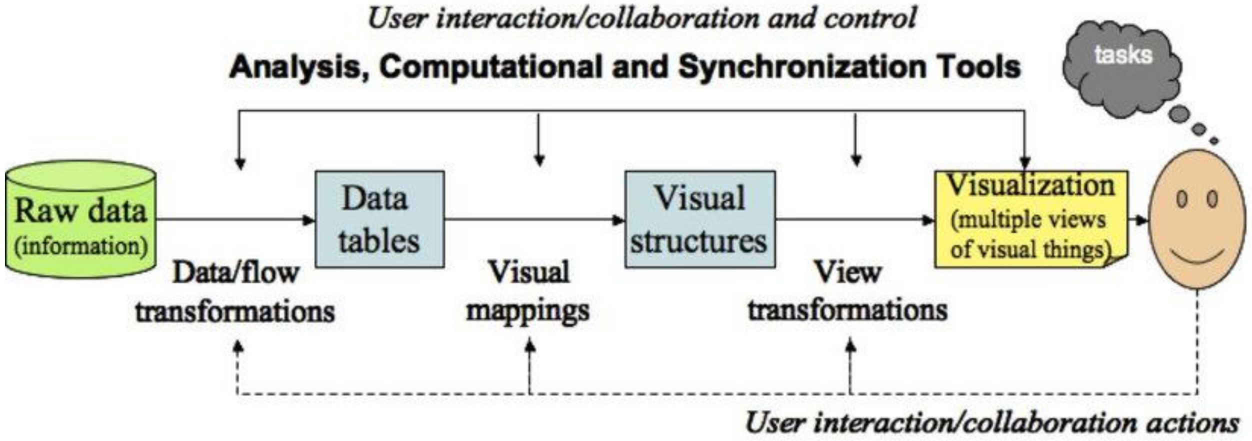




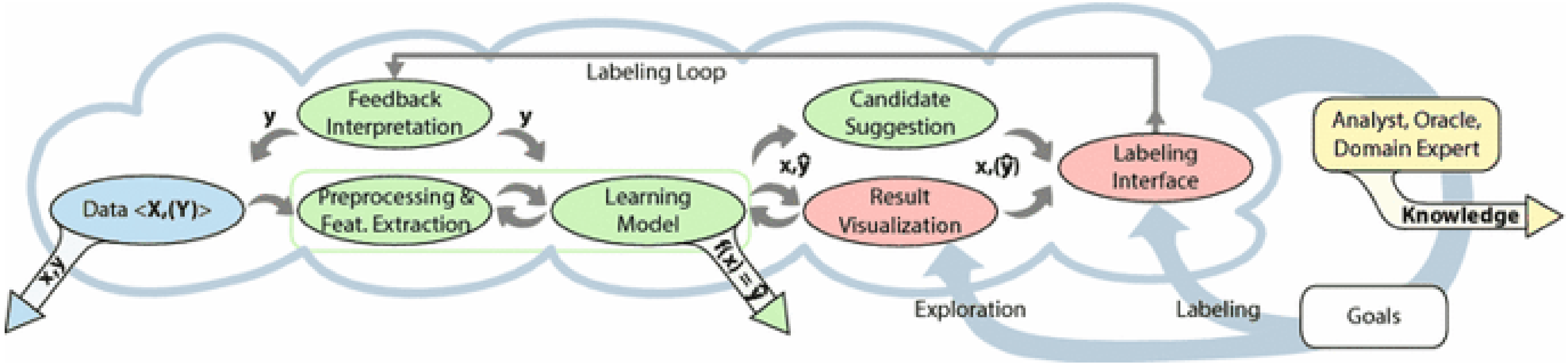


Explainable ML

- ML explanation != completely know how ML work



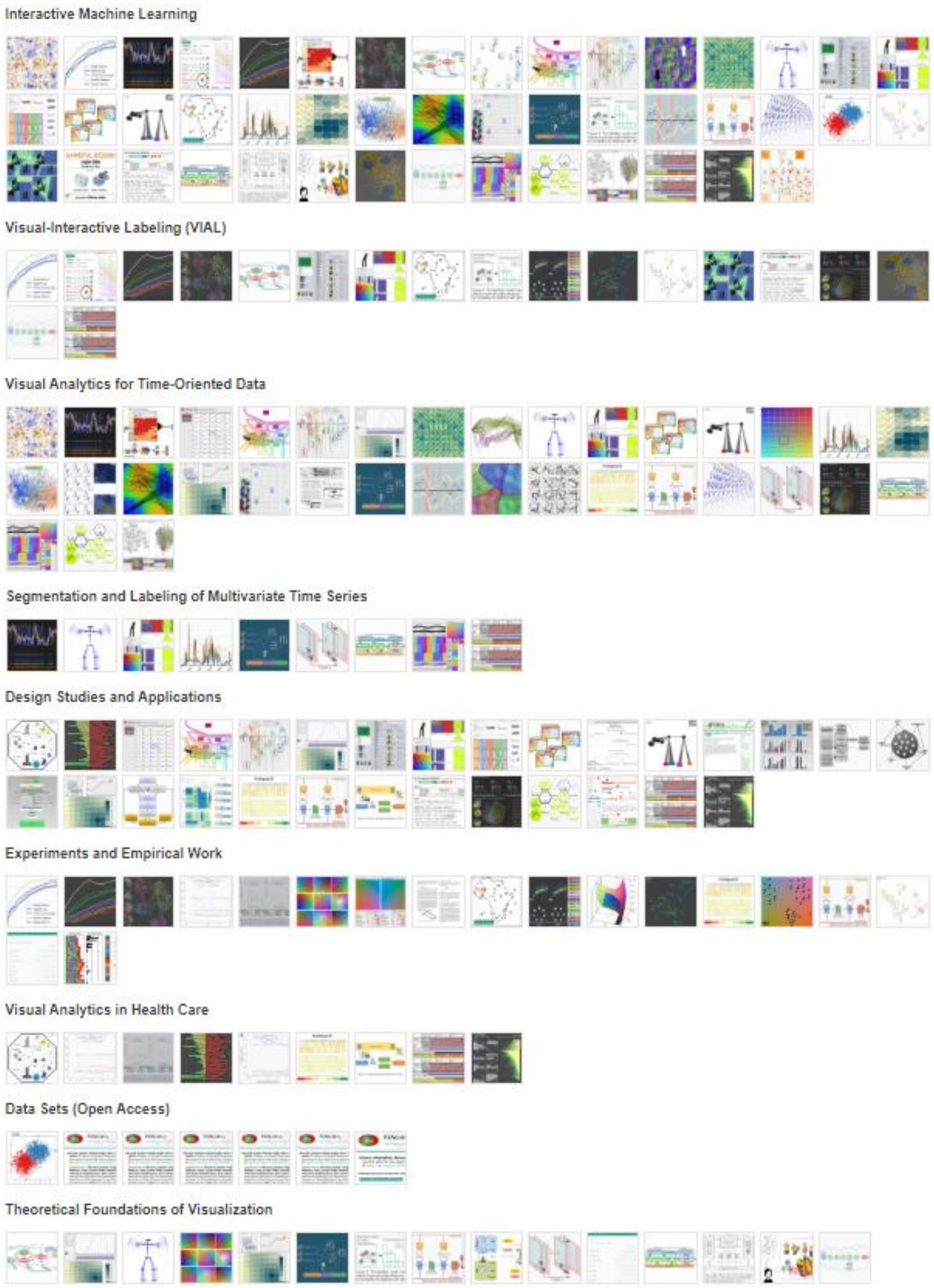
WARD et al. 2010



Bernard et al. 2018

Explainable ML

http://juergen-bernard.de/



Interactive Machine Learning

Visual-Interactive Labeling (VIAL)

Visual Analytics for Time-Oriented Data

Segmentation and Labeling of Multivariate Time Series

Design Studies and Applications

Experiments and Empirical Work

Visual Analytics in Health Care

Data Sets (Open Access)

Theoretical Foundations of Visualization