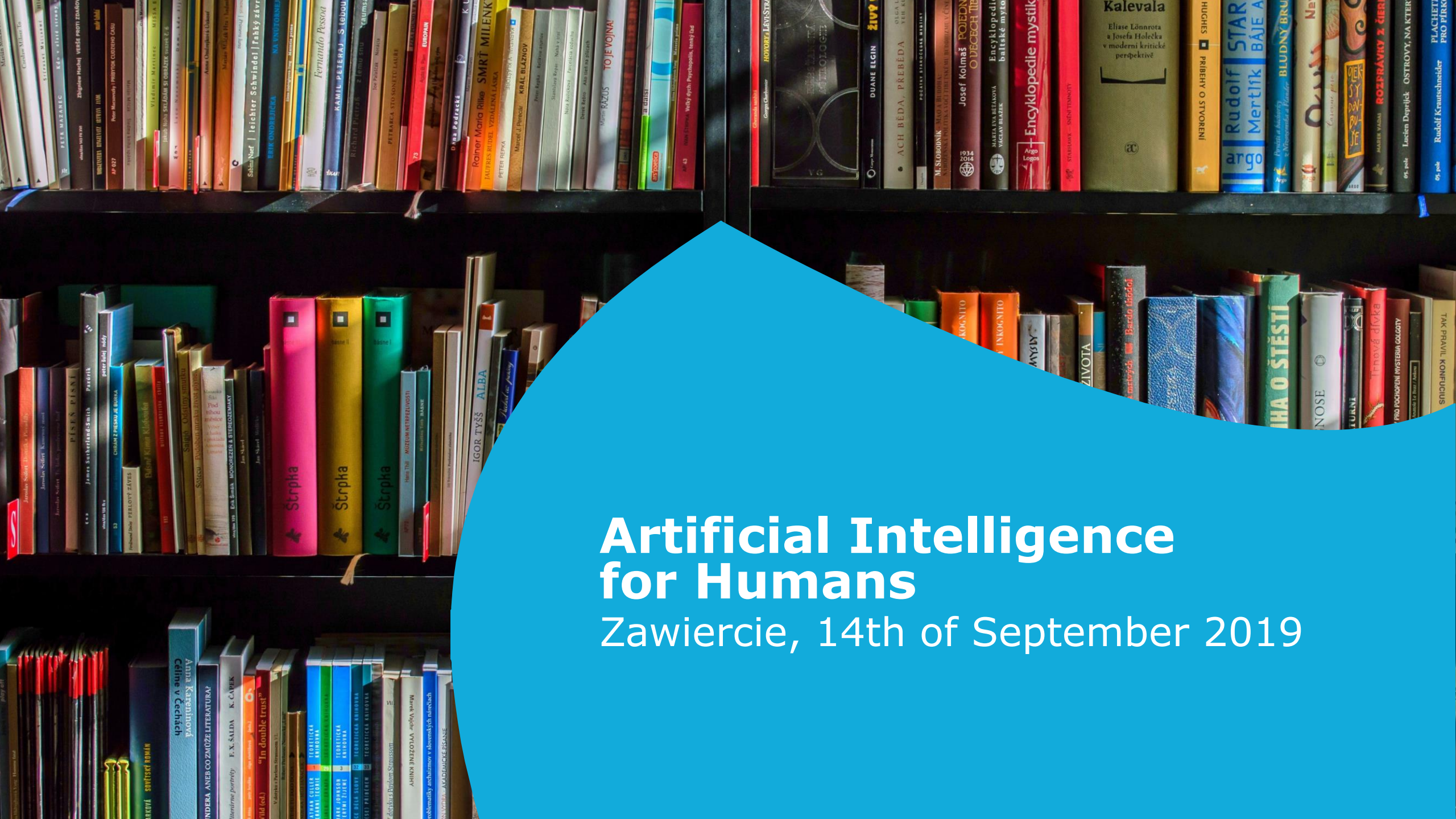




Artificial Intelligence for Humans

Zawiercie, 14th of September 2019

Expectation vs Reality

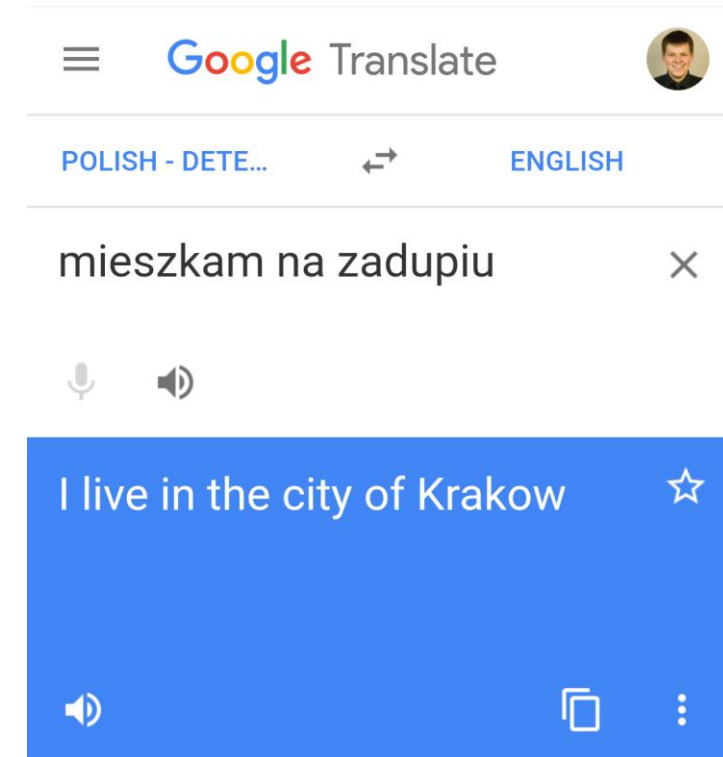


Expectation

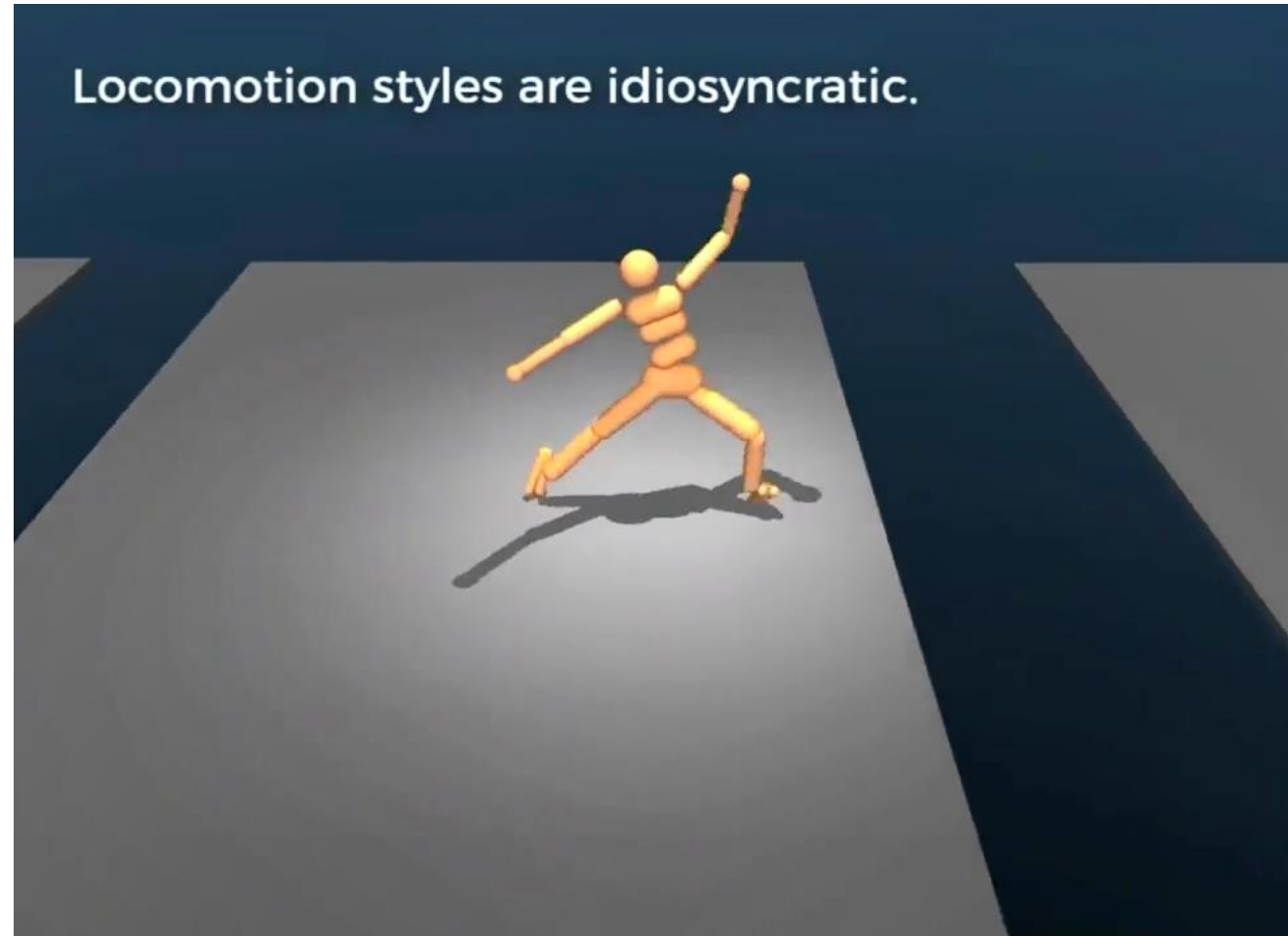


Source: bit.ly/2I6sSAN

Reality



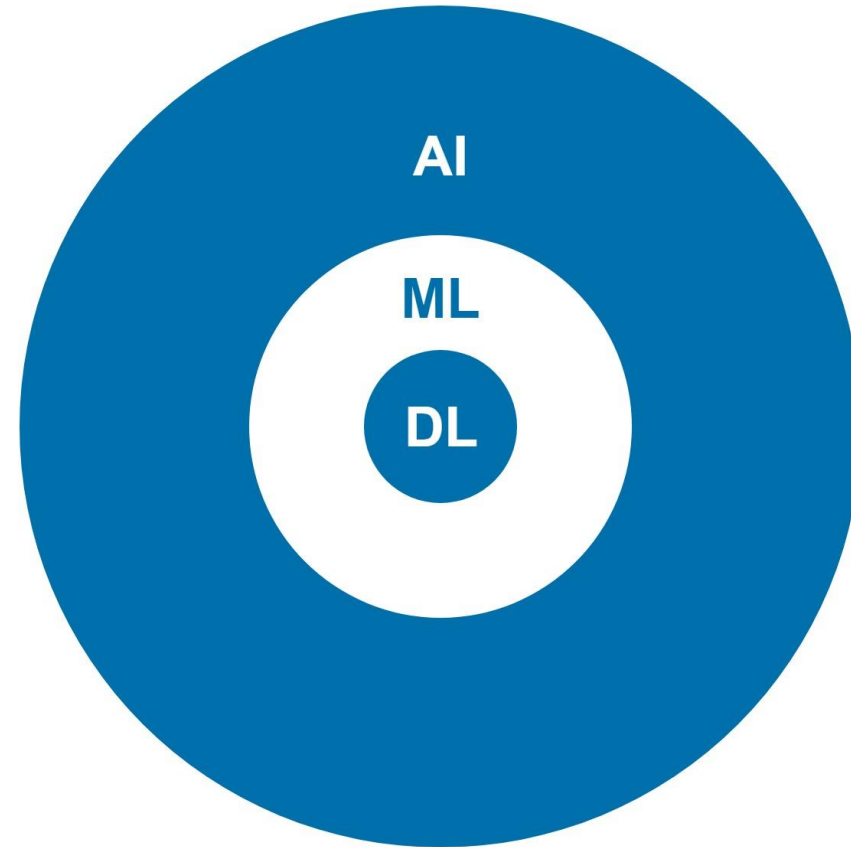
Expectation vs Reality



Source: bit.ly/2P0nqTi

Terminology

Artificial Intelligence vs Machine Learning



Artificial Intelligence vs Machine Learning



AI

Artificial Intelligence

Capability of a machine to imitate intelligent human behavior

ML

Machine Learning

Process used to achieve Artificial Intelligence

DL

Deep Learning

Algorithms that try to mimic the brain

Artificial Intelligence vs Machine Learning



Mat Velloso

@matvelloso



Difference between machine learning and AI:

If it is written in Python, it's probably machine learning

If it is written in PowerPoint, it's probably AI

♡ 23K 2:25 AM - Nov 23, 2018



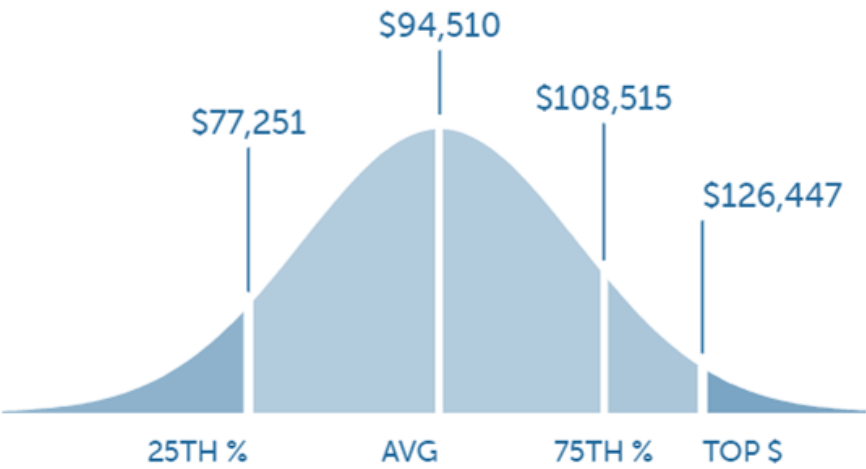
💬 8,467 people are talking about this



Artificial Intelligence vs Machine Learning



Compensation



Salaries for Machine Learning Engineers

TOTAL COMPENSATION	
25TH PERCENTILE	\$77,251
AVERAGE	\$94,510
75TH PERCENTILE	\$108,515
TOP EARNERS	\$126,447

BREAK DOWN: AVERAGE

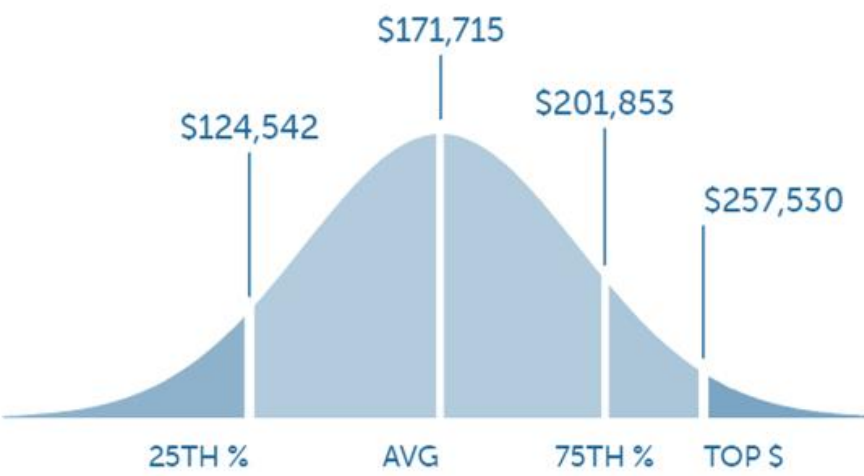
BASE	\$94,510
------	----------

Source: bit.ly/30I1Pq2

Artificial Intelligence vs Machine Learning



Compensation

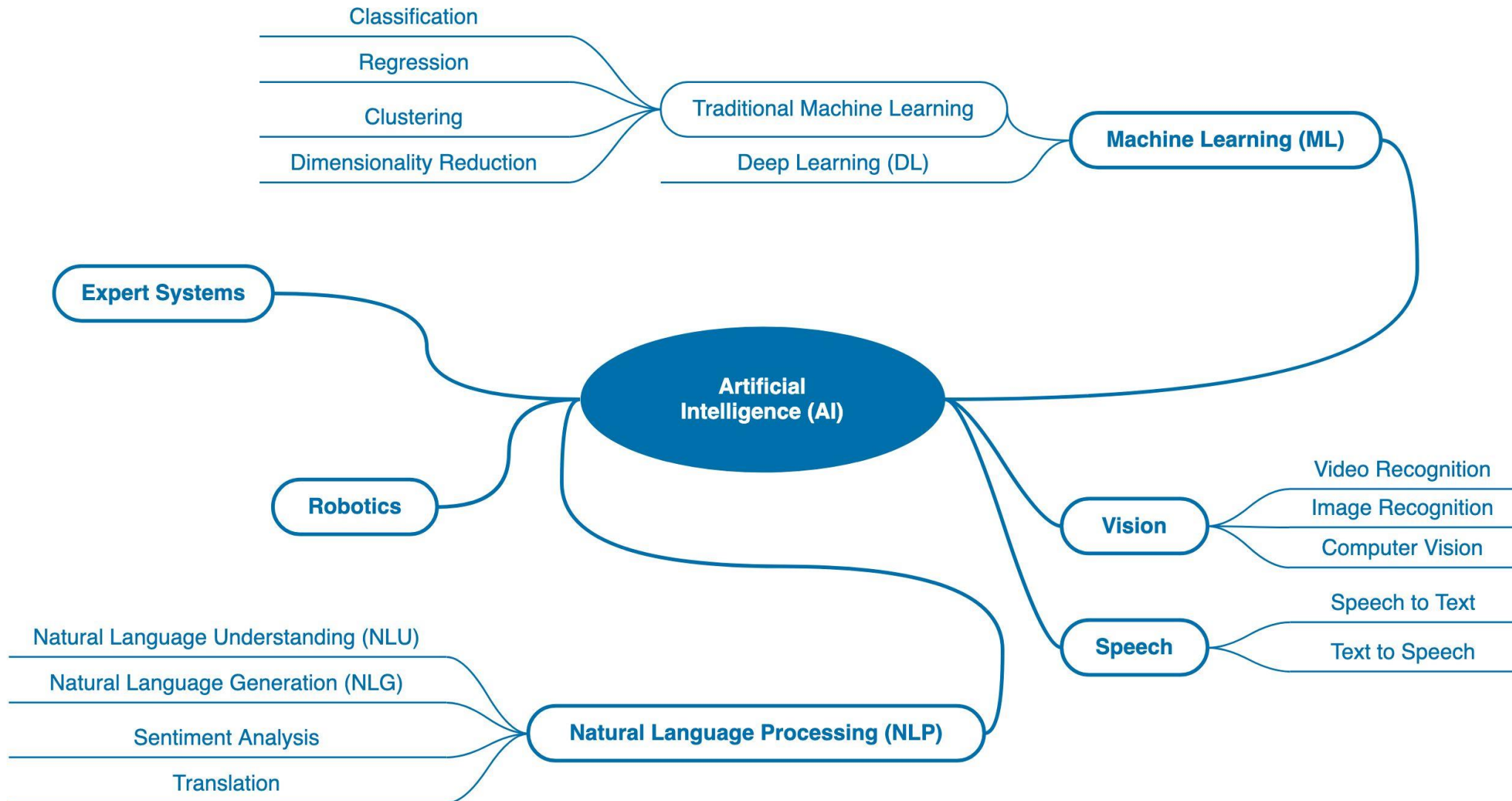


Salaries for Artificial Intelligence Engineers

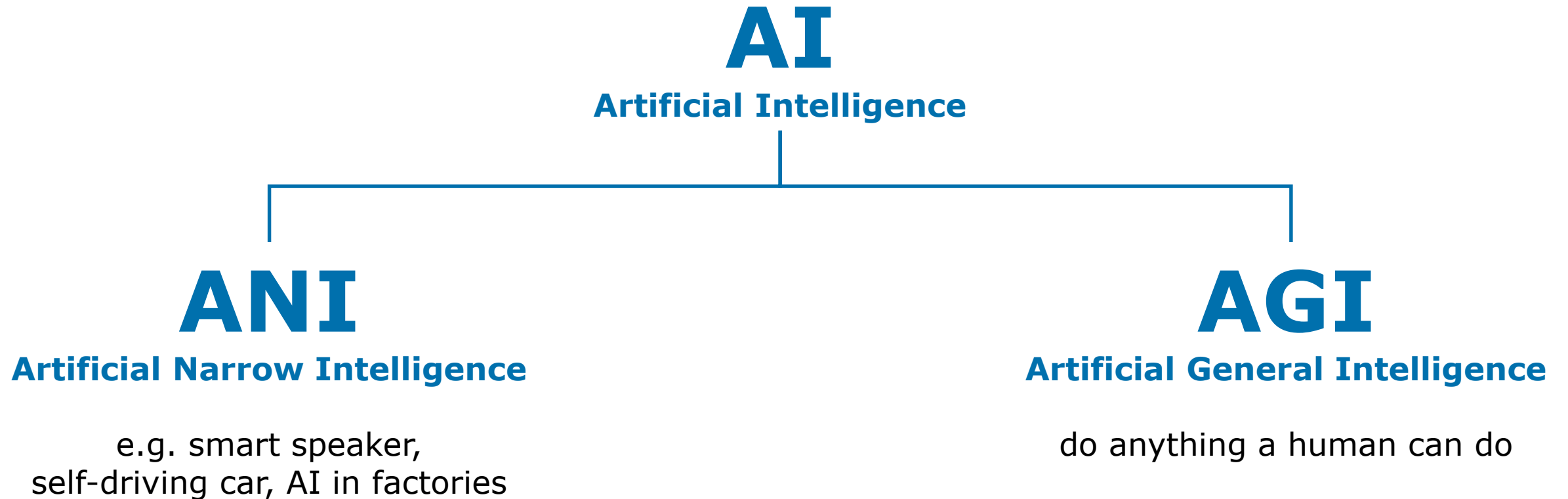
TOTAL COMPENSATION	
25TH PERCENTILE	\$124,542
AVERAGE	\$171,715
75TH PERCENTILE	\$201,853
TOP EARNERS	\$257,530

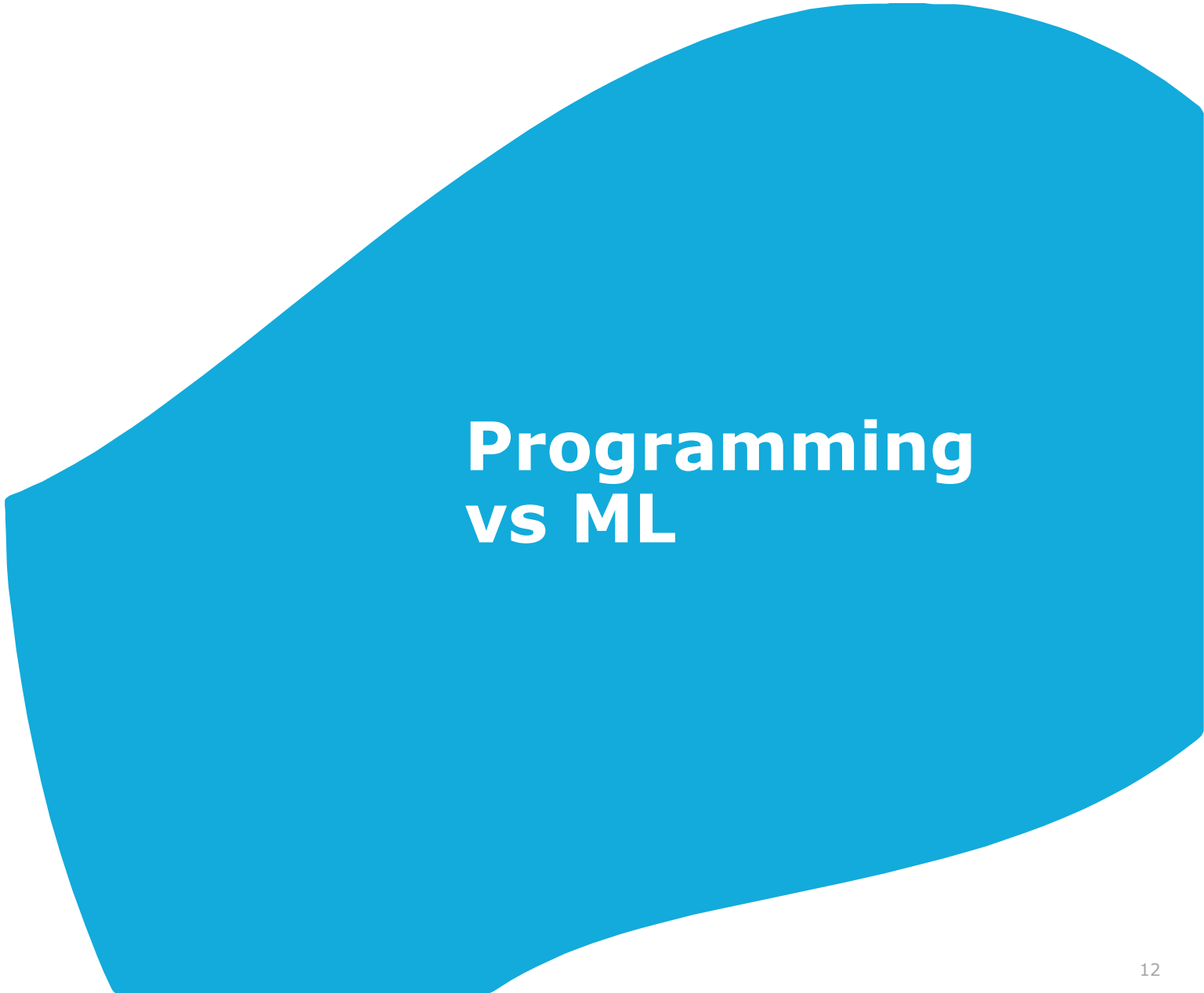
BREAK DOWN: AVERAGE	
BASE	\$114,857
EQUITY	\$39,666
BONUS	\$17,191
SIGNING BONUS	\$9,268

Source: bit.ly/2M9egHU



ANI vs AGI



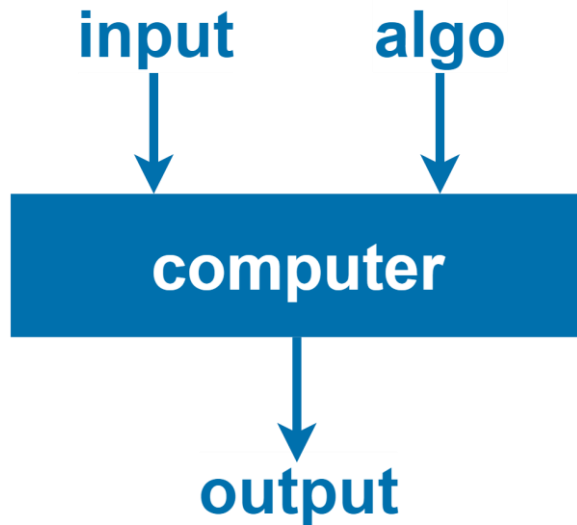


Programming vs ML

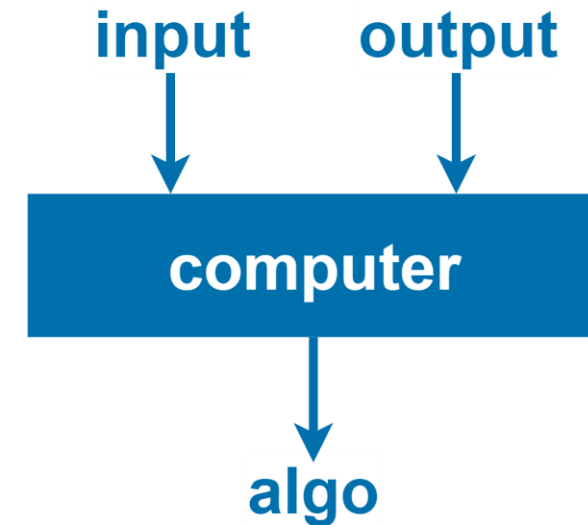
Traditional Programming vs Machine Learning



Traditional Programming



Machine Learning



Traditional Programming vs Machine Learning



Traditional Programming

$$1 + 1 = \square$$

$$1 + 2 = \square$$

Machine Learning

$$1 \square 1 = 2$$

$$1 \square 2 = 3$$

Traditional Programming vs Machine Learning



Traditional Programming

$$1 + 1 = \square$$

$$1 + 2 = \square$$

$$1 + 3 = 4$$

Machine Learning

$$1 \square 1 = 2$$

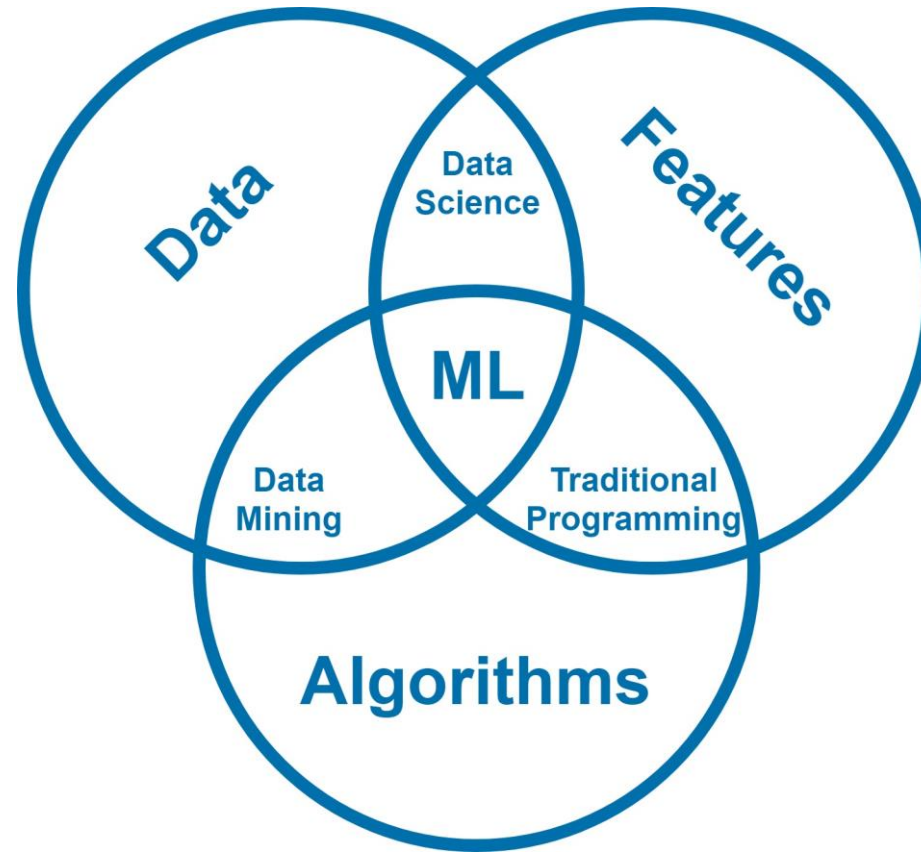
$$1 \square 2 = 3$$

$$0 + 1*1 + 1*3 = 4$$

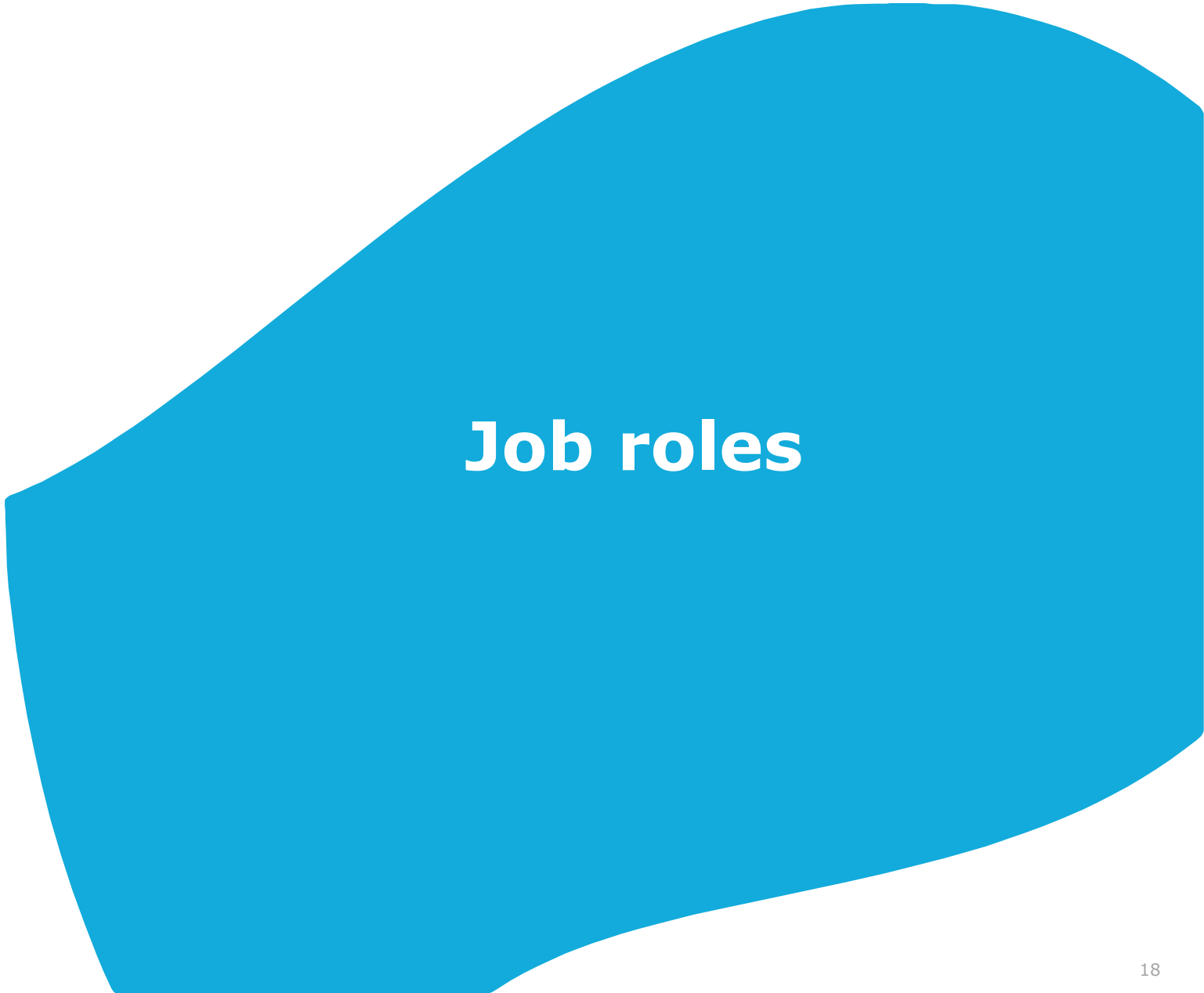


Data Science vs Data Mining

Data Science vs Data Mining



Source: bit.ly/2kFi4Uz



Job roles

Data science job roles

Part 1



Data Scientist

- Python
- R
- Pandas, Apache Spark
- Plotly
- Scikit-learn
- TensorFlow
- Jupyter
- PowerPoint



Artificial Intelligence Engineer

- Python
- Cloud computing
- Scikit-learn
- TensorFlow
- Jupyter
- PowerPoint
- Draw.io



Data Architect

- Cloud computing
- Data warehouse
- PowerPoint
- Draw.io
- Excel



Data science job roles

Part 2



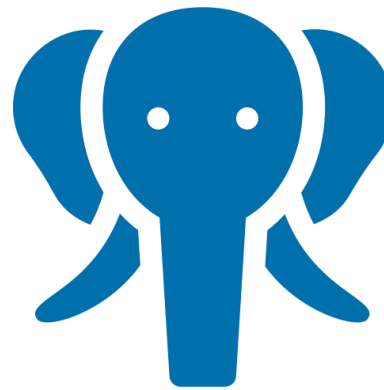
Data Engineer

- Python
- Pandas, Apache Spark
- Apache Airflow, Luigi
- Cloud computing
- Linux
- Hadoop
- Git



Big Data Specialist

- Java
- Hadoop
- Linux
- Git



Data Analyst

- BI tools
- SQL
- Excel
- PowerPoint





ML tasks

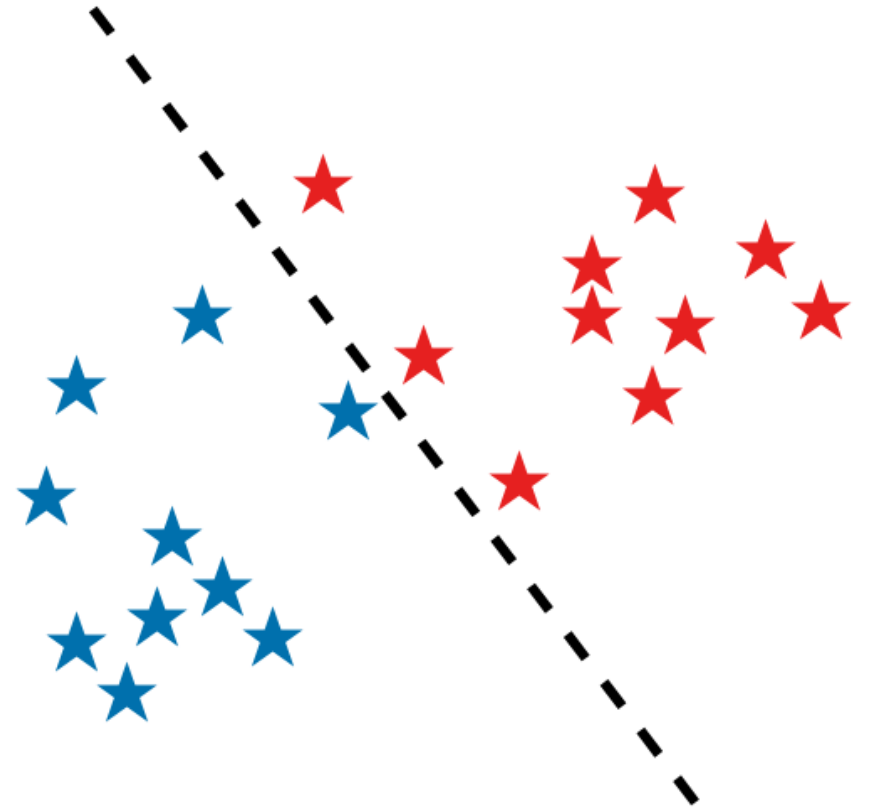
Classification



Prediction value as category like yes/no, positive/negative, etc.

Examples:

- Email filtering.
- Credit scoring.
- Finding named entities in text: person, organization, location, etc.
- Language identification.
- Prediction of diabetes.



Credit scoring



Classes: *1* – grant credit, *0* – not grant credit.

Features:

1. age
2. civil status
3. credit limit
4. education
5. gender
6. income
7. job title
8. work experience



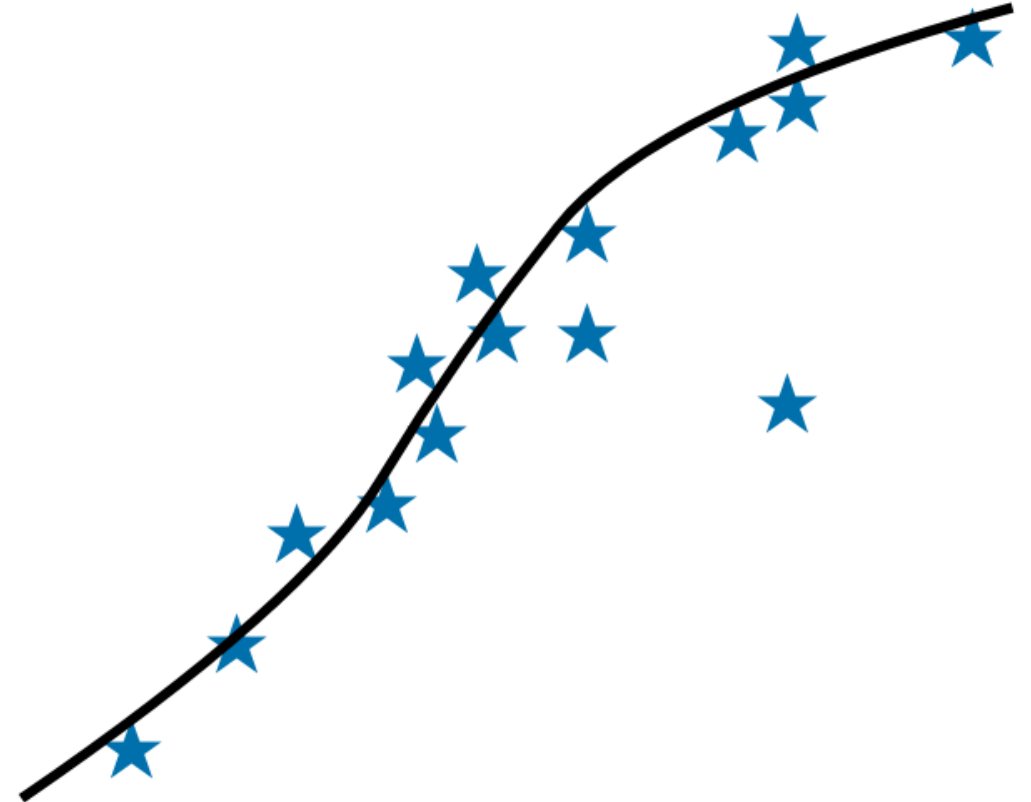
Regression



Prediction value as continuous value.

Examples:

- Company's yearly revenue prediction.
- Foreign exchange market prediction.
- Job salary prediction.
- Predicting house and car price.
- Stock market prediction.



Stock market prediction



Output: daily price of share (close price).

Features:

1. weekday
2. previous open price
3. previous open price for high correlated share
4. previous low price
5. previous high price
6. previous volume
7. previous stock market index
8. previous exchange rate



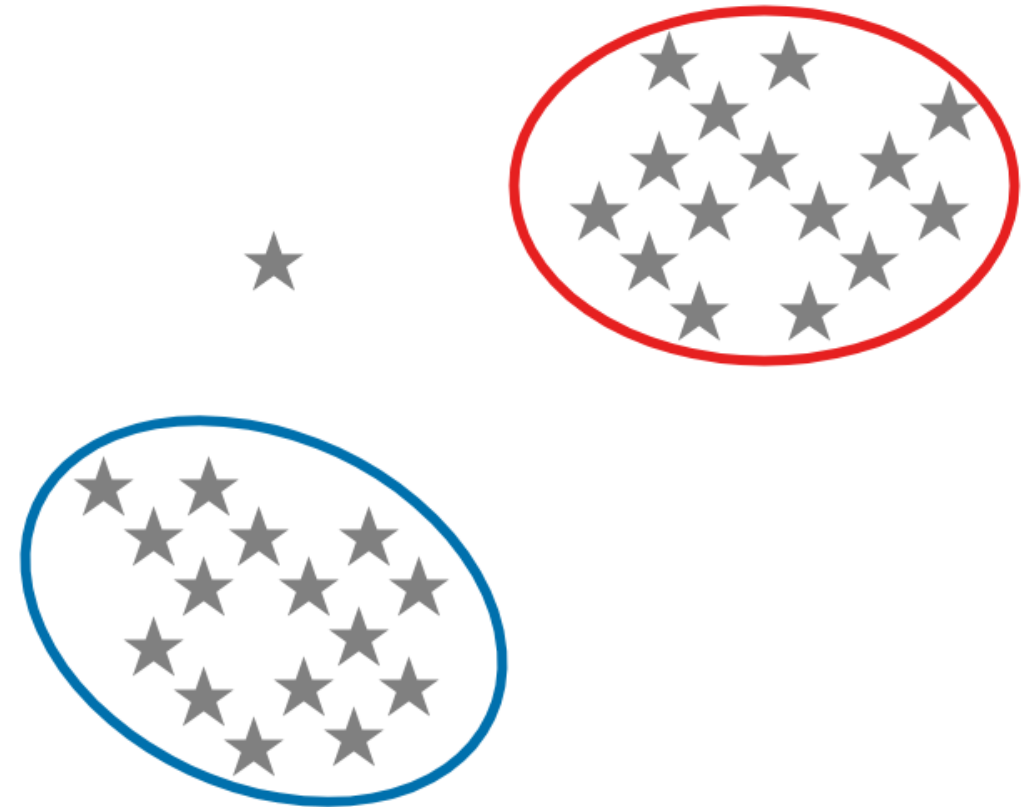
Clustering



Grouping a set of points to clusters.

Examples:

- Fraudulent bank transaction identification.
- Human activity recognition using smartphones.
- Image compression.
- Recommender system.
- One-class classification.



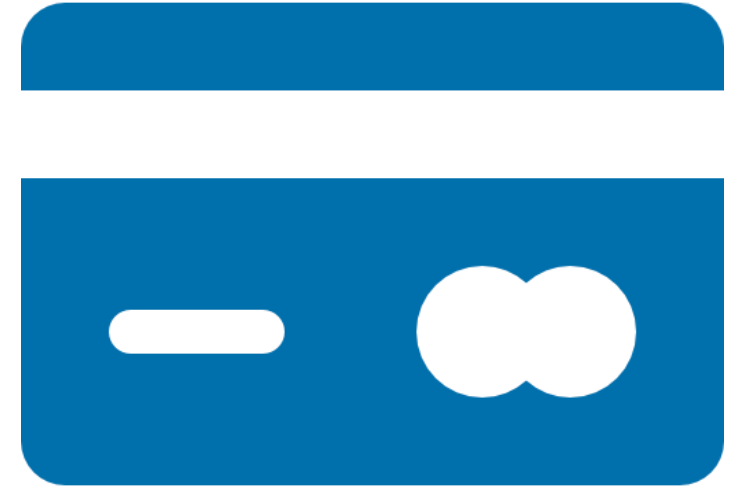
Fraudulent bank transaction identification



Output: money transfer approve or reject.

Features:

1. amount
2. currency in
3. currency out
4. quarter
5. receiver
6. receiver country
7. sender
8. sender country
9. weekday



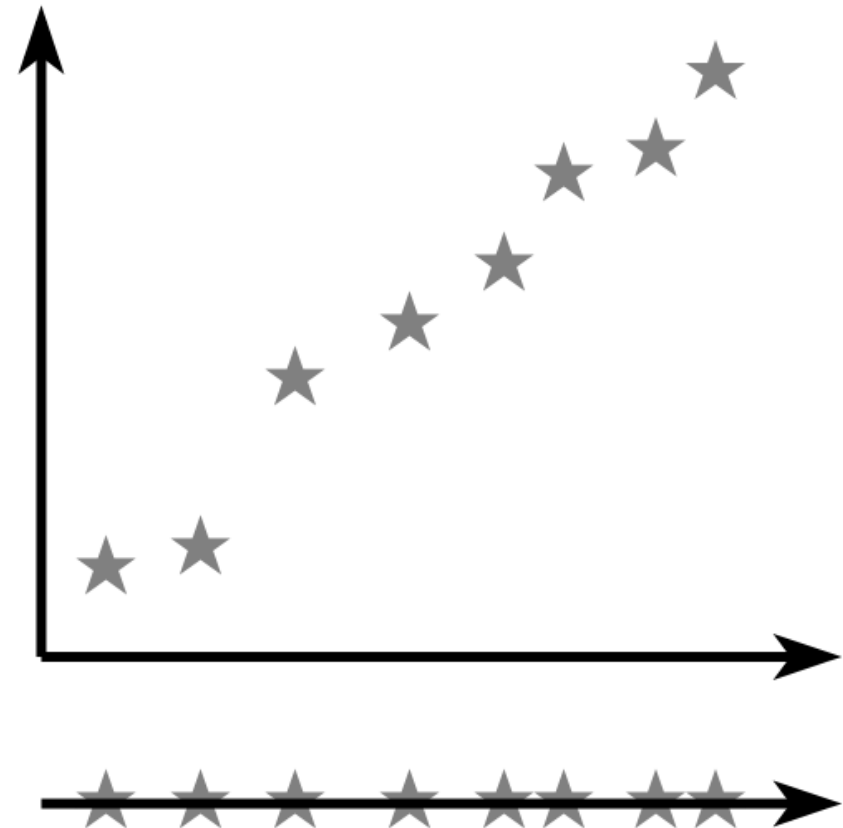


Dimensionality reduction

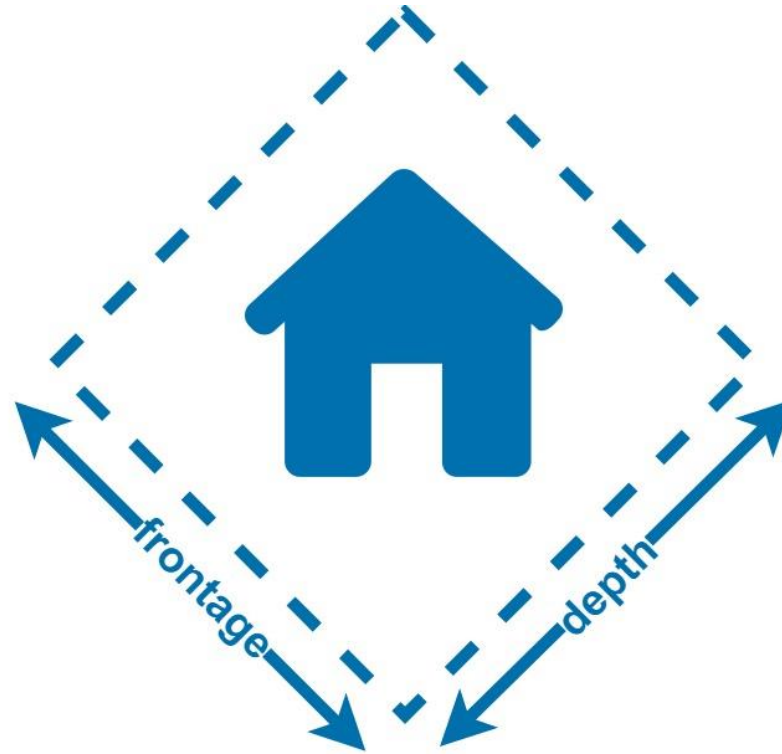
Reducing the number of features.

Examples:

- Data visualization.
- Fake image detection.
- Recommender system.
- Risk management.
- Topic Modelling.



Housing prices prediction



$$\text{area} = \text{frontage} * \text{depth}$$

Recommender system



	movie1	movie2	...	movieN
name1				
name2				
...				
nameN				

name1				
name2				
...				
nameN				

Recommender system



	movie1	movie2	...	movieN
name1				
name2				
...				
nameN				

	genre1	genre2	...	genreN
name1				
name2				
...				
nameN				

BREAK

Test

Part 1



- Anomaly detection.
- Cancer prediction.
- Champions League winner prediction.
- Character recognition.
- E-commerce business forecasting.
- Employee attrition prediction.
- Employee's unplanned absence prediction.
- Image clustering.

Test

Part 2



- IT system availability prediction.
- Joiner prediction in recruitment.
- Map points combining.
- Market basket analysis.
- Market segmentation (types of customers, loyalty).
- Project overrun prediction.
- Sentiment analysis.
- Telecom industry churners prediction.

Use cases

Use cases



Manufacturing

- Predictive maintenance
- Warranty reserve estimation
- Propensity to buy
- Demand forecasting
- Process optimization
- Telematics

Retail

- Predictive inventory planning
- Recommendation engines
- Upsell and cross-channel marketing
- Market segmentation and targeting
- Customer ROI and LTV

Healthcare and life sciences

- Real-time alerts and diagnostics
- Disease identification and risk stratification
- Patient triage optimization
- Proactive health management
- Healthcare provider sentiment analysis

Financial services

- Risk analytics and regulation
- Credit worthiness evaluation
- Customer segmentation
- Cross-selling and upselling
- Sales and marketing campaign management

Travel and hospitality

- Aircraft scheduling
- Dynamic pricing
- Feedback and interaction analysis
- Customer complaint resolution
- Traffic patterns and congestion management

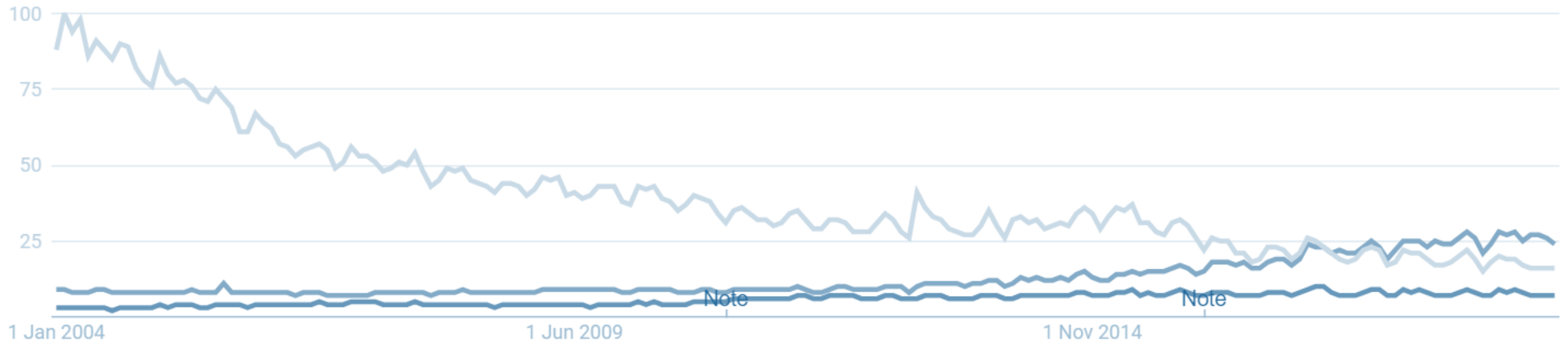
Energy, feedstock and utilities

- Power usage analytics
- Seismic data processing
- Carbon emissions and trading
- Customer-specific pricing
- Smart grid management
- Energy demand and supply optimization

Source: bit.ly/2INPdh6

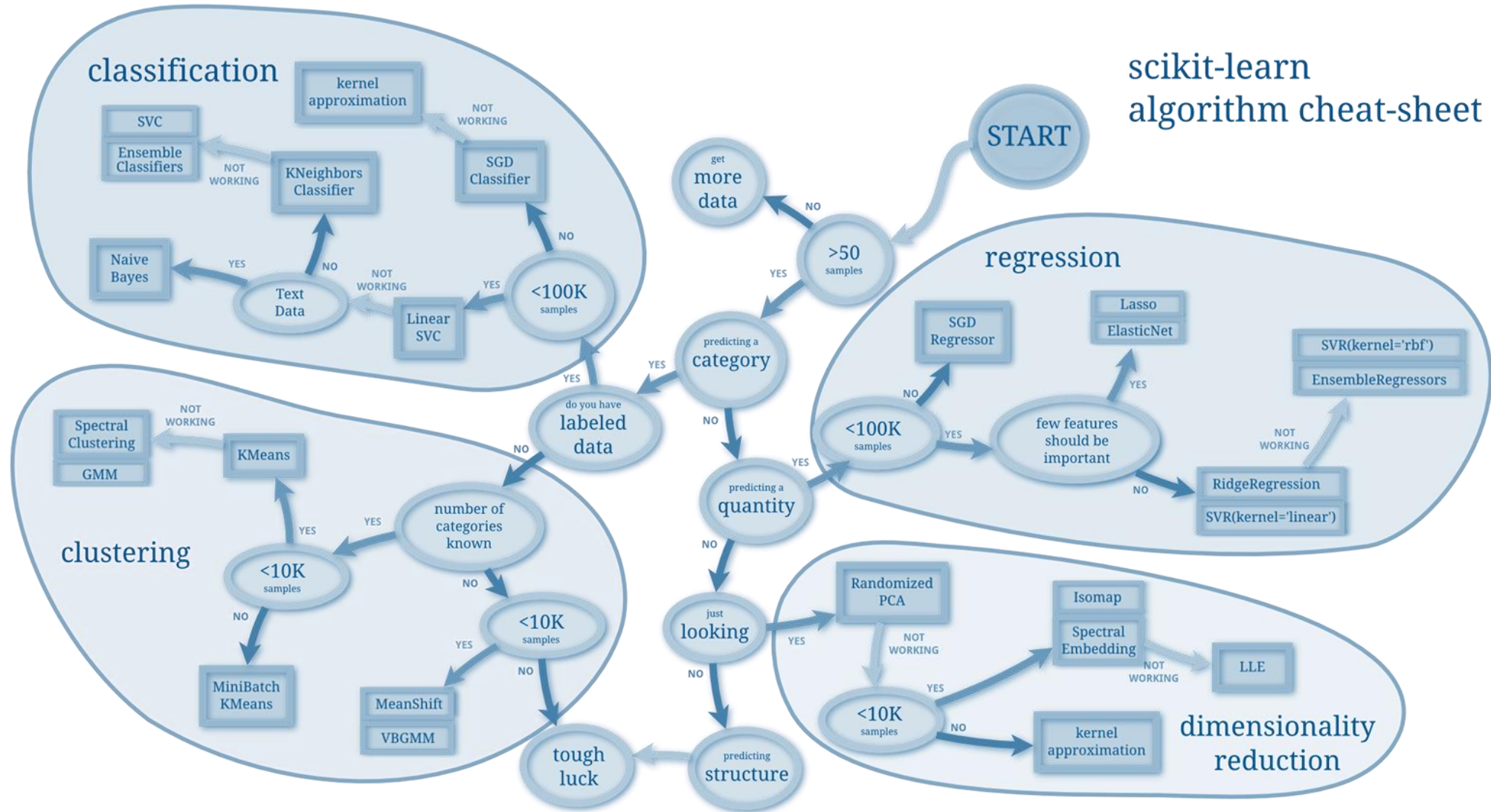
Python vs R

Python vs R vs Java



Source: bit.ly/2Ojajhs

Algorithms



Source: bit.ly/2HuB6II



Let's start

Classification in Python

Predict Diabetes



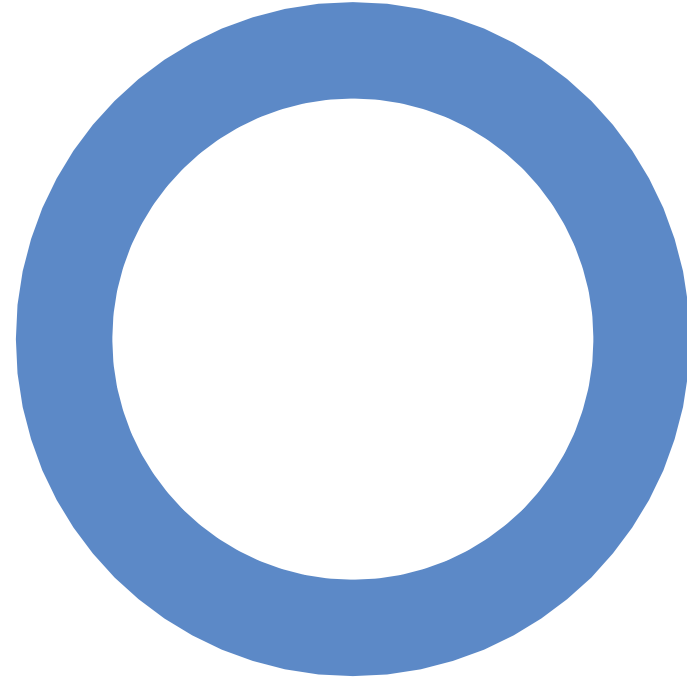
Classes: *1* – diabetic, *0* – non-diabetic.

Features:

1. number of pregnancies
2. plasma glucose concentration
3. blood pressure
4. skin fold thickness
5. serum insulin
6. body mass index
7. diabetes pedigree function
8. age

Data and code source: bit.ly/2Mbo53w

Image source: bit.ly/2MN4Myi



Classification in R

Banknote Auth



Classes: 1 – genuine, 0 – counterfeit (fake).

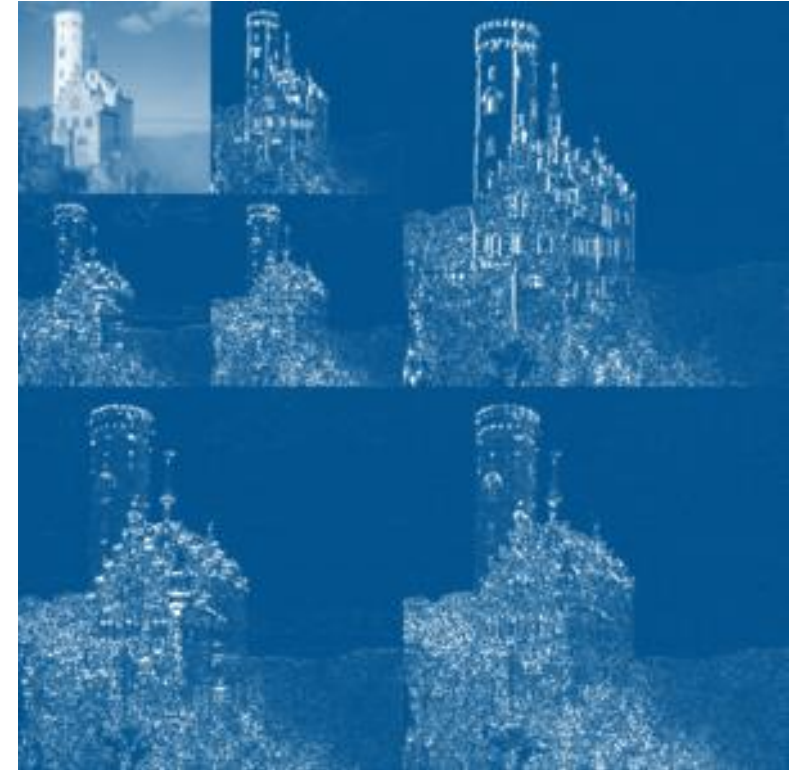
Features:

1. [variance](#) of Wavelet Transformed image
2. [skewness](#) of Wavelet Transformed image
3. [kurtosis](#) of Wavelet Transformed image
4. entropy of image

Data source: bit.ly/2JpLsp3

Code source: bit.ly/2K2wcP3

Image source: bit.ly/2kWjwyw



jota.smcebi.us.edu.pl

bit.ly/DEMOPYCON

Tips and Tricks

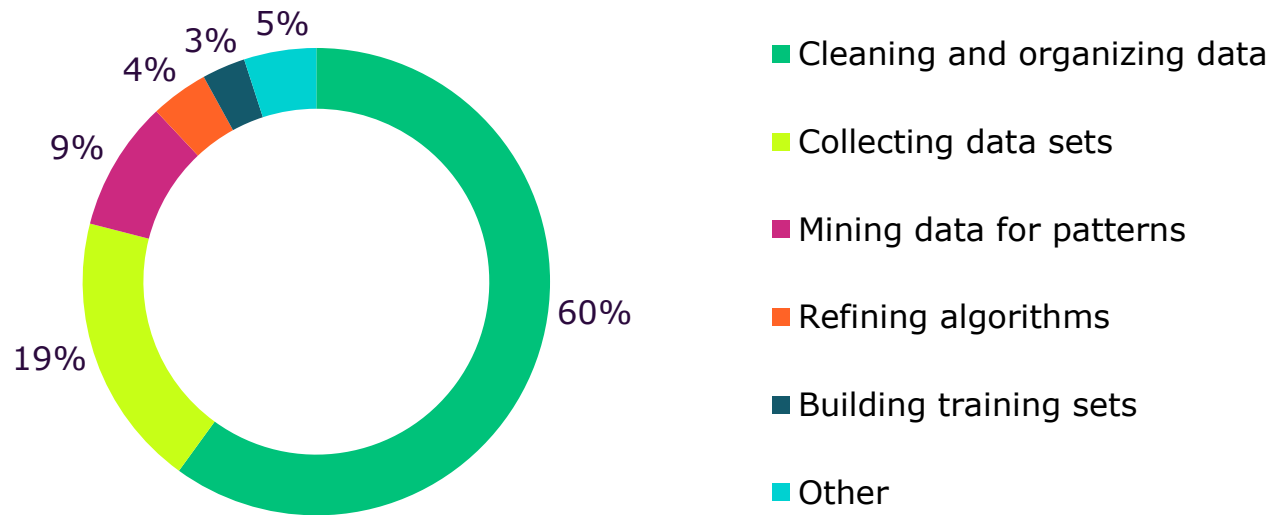
by badly-behaved Data Scientist

1/5 Data preparation



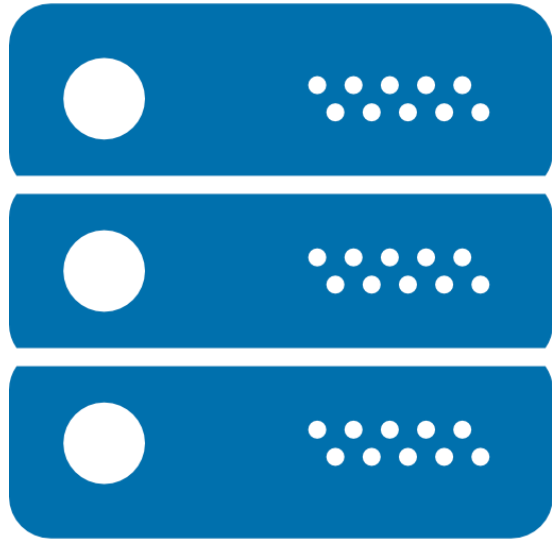
**90% data
preparation
10%
analysis**

1/5 Data preparation



Source: bit.ly/2kHWM8L

**Data
preparation
accounts for
about 80% of
the work**



**# of
transistors
x2 every
two years**

3/5 Domain knowledge



**Domain
knowledge
is important**

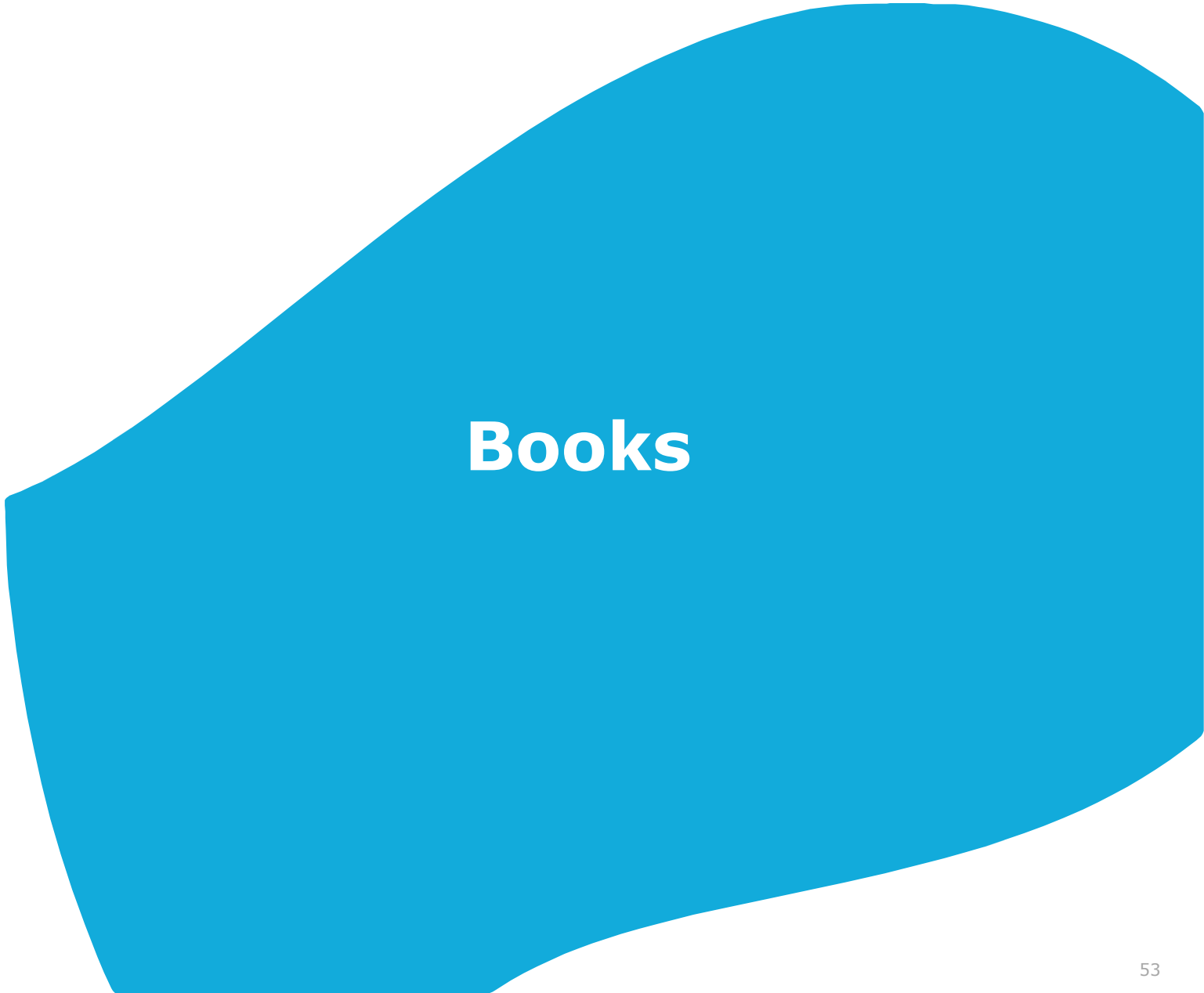
4/5 Magic numbers



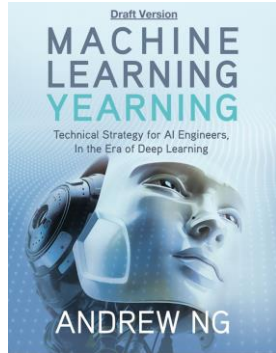
70% / 30%
train / test



**Polish is the
hardest
language!**

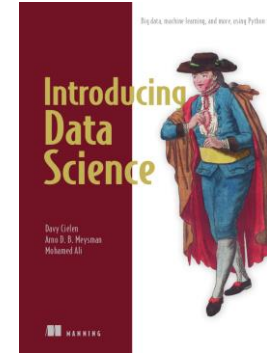


Books



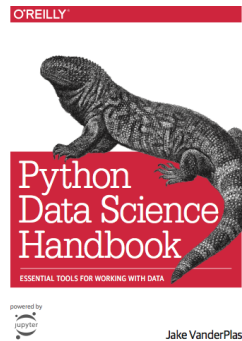
Machine Learning Yearning by Andrew Ng

- [in English](#)
- [in Russian](#)



Introducing Data Science by Davy Cielen, Arno Meysman, and Mohamed Ali

- in English
- in Russian



Python Data Science Handbook by Jake VanderPlas

- [in English](#)
- in Russian



AI Leaders

AI Leaders



Geoffrey Hinton

- 1947, UK
- [Wikipedia](#)
- [Website](#)
- **Twitter:** geoffreyhinton
- geoffrey.hinton@gmail.com



Yann LeCun

- 1960, France
- [Wikipedia](#)
- [Website](#)
- **Twitter:** ylecun
- **Facebook:** yann.lecun
- **LinkedIn:** yann-lecun-0b999
- yann@fb.com



Yoshua Bengio

- 1964, France
- [Wikipedia](#)
- [Website](#)
- **Facebook:** yoshua.bengio
- **LinkedIn:** yoshuabengio
- yoshua.bengio@umontreal.ca

AI Leaders



Andrew Ng

- 1976, UK
- [Wikipedia](#)
- [Website](#)
- **Twitter:** AndrewYNg
- **Facebook:** andrew.ng.96
- **LinkedIn:** andrewyng
- ang@cs.stanford.edu



Fei-Fei Li

- 1976, Beijing
- [Wikipedia](#)
- **Twitter:** drfeifei
- **Facebook:** feifei.li.372
- **LinkedIn:** fei-fei-li-4541247
- feifeili@cs.stanford.edu



Siraj Raval

- 1991, US
- [YouTube](#)
- [Website](#)
- **Twitter:** sirajraval
- **Facebook:** sirajologyyy
- **LinkedIn:** sirajraval
- hello@sirajraval.com

About Capgemini

A global leader in consulting, technology services and digital transformation, Capgemini is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms. Building on its strong 50-year heritage and deep industry-specific expertise, Capgemini enables organizations to realize their business ambitions through an array of services from strategy to operations. Capgemini is driven by the conviction that the business value of technology comes from and through people. It is a multicultural company of over 200,000 team members in more than 40 countries. The Group reported 2018 global revenues of EUR 13.2 billion.

Learn more about us at

www.capgemini.com



People matter, results count.

This presentation contains information that may be privileged or confidential and is the property of the Capgemini Group.
Copyright © 2019 Capgemini. All rights reserved.

Ruslan Korniiichuk

Senior Consultant

Capgemini

ruslan.korniichuk@capgemini.com

+48 791 065 877