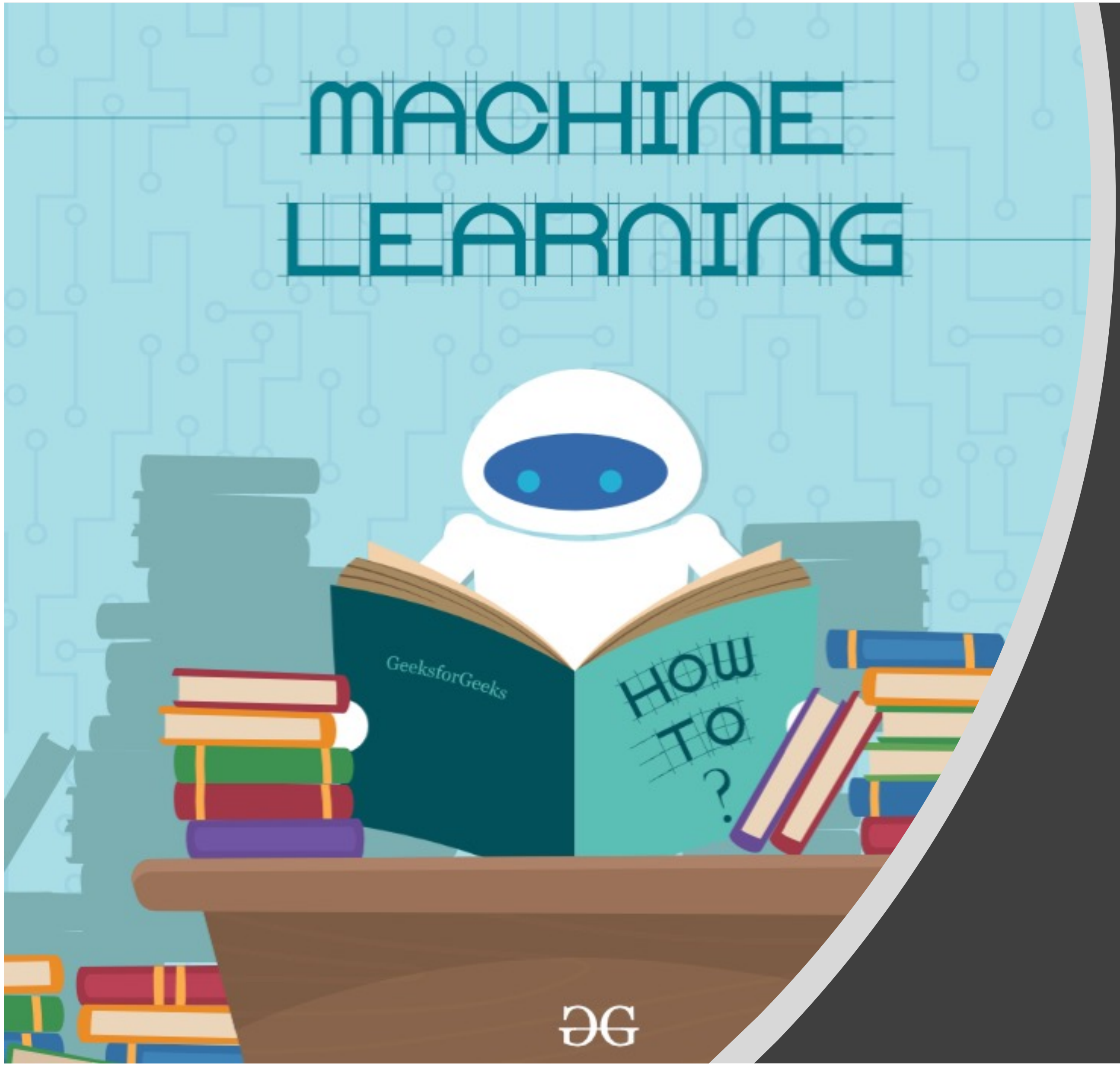


Machine Learning for Economics

Marina Rizzi

18 July 2022

"Labor Economics" Course by Francesco Devicienti

Overview of this lecture:

-What it is Machine Learning (and some common models used there)

-How do Economists use Machine Learning?

-Let's dive into some coding!

Aim of this lecture:

-Give an introductive view of what Machine learning is, and introduce to the main methods and terminology used

-Give examples of where Machine learning could be used in Economics

-Give a glimpse of the implementation of some codes and algorithms in Python

Let's start with a question: what is Artificial Intelligence, in your opinion?

Ansa

L'intelligenza artificiale per decifrare una lingua sconosciuta

Utilizzare l'intelligenza artificiale per aiutare a decifrare una lingua antica che ancora oggi resta sconosciuta.

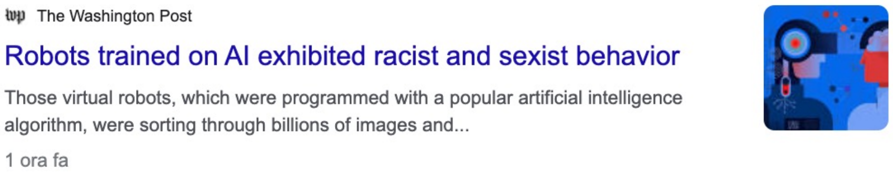
23 ore fa

greenMe

L'intelligenza artificiale ci sta davvero aiutando a decifrare una lingua di 3.500 anni fa

Un gruppo di ricerca dell'Università di Bologna ha usato l'intelligenza artificiale per decifrare una lingua sconosciuta, il cipro-minoico.

2 ore fa

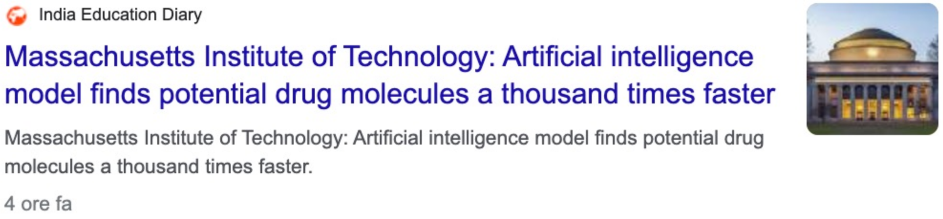


The Washington Post

Robots trained on AI exhibited racist and sexist behavior

Those virtual robots, which were programmed with a popular artificial intelligence algorithm, were sorting through billions of images and...

1 ora fa



India Education Diary

Massachusetts Institute of Technology: Artificial intelligence model finds potential drug molecules a thousand times faster

Massachusetts Institute of Technology: Artificial intelligence model finds potential drug molecules a thousand times faster.

4 ore fa

Tgcom24

Fake news, Zuckerberg arruola l'Intelligenza Artificiale anti-bufale

Meta, la società madre di Facebook e Instagram di Mark Zuckerberg, sta testando Sphere, uno strumento basato sull'Intelligenza Artificiale...

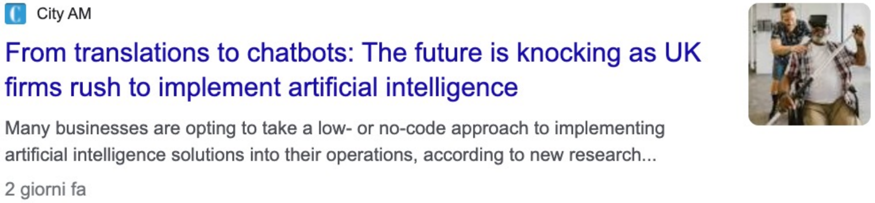
23 ore fa

Forbes

Artificial Intelligence: Not A Panacea For Supply Chain Issues, But Extremely Helpful

Artificial intelligence is another area of technology investment that holds potential, and early results are promising.

1 giorno fa

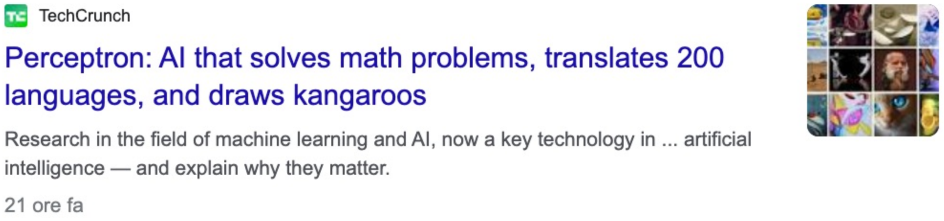


City AM

From translations to chatbots: The future is knocking as UK firms rush to implement artificial intelligence

Many businesses are opting to take a low- or no-code approach to implementing artificial intelligence solutions into their operations, according to new research...

2 giorni fa



TechCrunch

Perceptron: AI that solves math problems, translates 200 languages, and draws kangaroos

Research in the field of machine learning and AI, now a key technology in ... artificial intelligence — and explain why they matter.

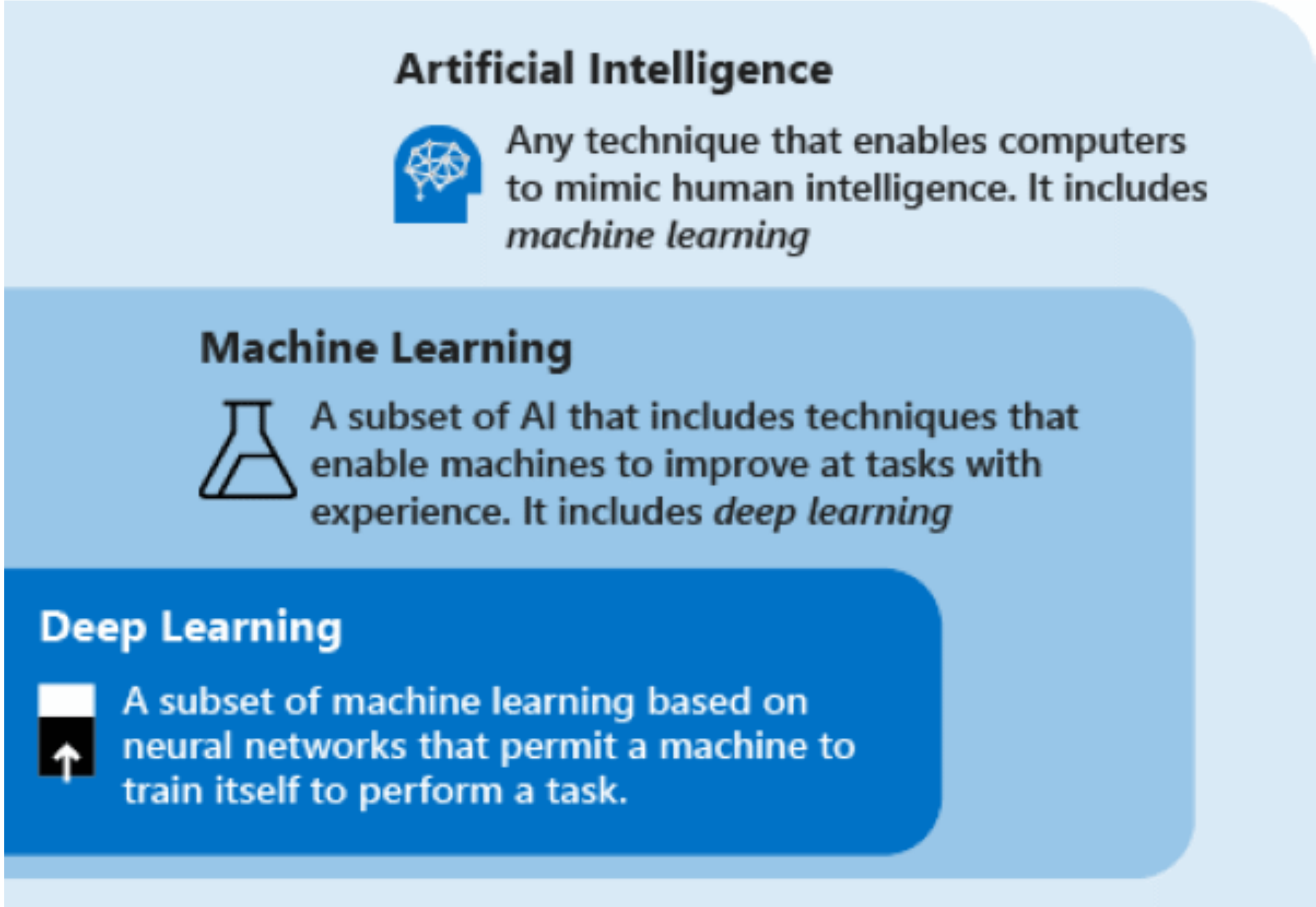
21 ore fa

Artificial Intelligence: (many definitions):

- ***"It is the study of how to train the computers so that computers can do things which at present human can do better."***

------> we want to add all the capabilities to a machine, that the human intelligence contains.

Machine Learning: often defined as a subfield of Artificial Intelligence (application of AI).



Artificial Intelligence

Any technique that enables computers to mimic human intelligence. It includes *machine learning*

Machine Learning

A subset of AI that includes techniques that enable machines to improve at tasks with experience. It includes *deep learning*

Deep Learning

A subset of machine learning based on neural networks that permit a machine to train itself to perform a task.

What is Machine Learning?

- Often categorized as a subfield of Artificial Intelligence
- Concretely: building *mathematical model* to help understand data.

- "Learning": when we give to these models ***tunable parameters*** that can be adapted to observed data.

-------> Once these models have been fit to previously seen data, they can be used to predict and understand aspects of newly observed data.

- Learning without being explicitely coded (but, taking data as input, it is able to generate decisions or prediction as outputs).

More technical definition of Machine Learning:

“Machine Learning is said to learn from experience E w.r.t some class of task T and a performance measure P if learners performance at the task T as measured by P improves with experience E.”

Applications developed thanks to Machine Learning:

Virtual Personal Assistants, Product Recommendations, Self Driving Cars, Translators Softwares, Fraud Detection, Preventinve Health Care

- Recent surge in the popularity of Machine Learning in a different variety of fields:

  - huge availability of data

  - more powerful computers

Machine Learning was important to solve a variety of problems that were difficult to be solved with "hard coding" (i.e. describing in a detailed way the action that the algorithm should have performed ----> Machine learning algorithms do not need a precise definition of what to do (i.e. how a chair looks like) but learn from the data the actions or the classifications they should perform)

- **Categories of Machine Learning Algorithms:**

-Supervised VS Unsupervised Algorithms (and also… Semi-supervised)
(one of the most important division)

-Supervised: classification or regression
-Unsupervised: clustering, dimensionality reduction, ….

-Reinforcement Learning

Supervised VS Unsupervised Machine Learning

Supervised Machine Learning

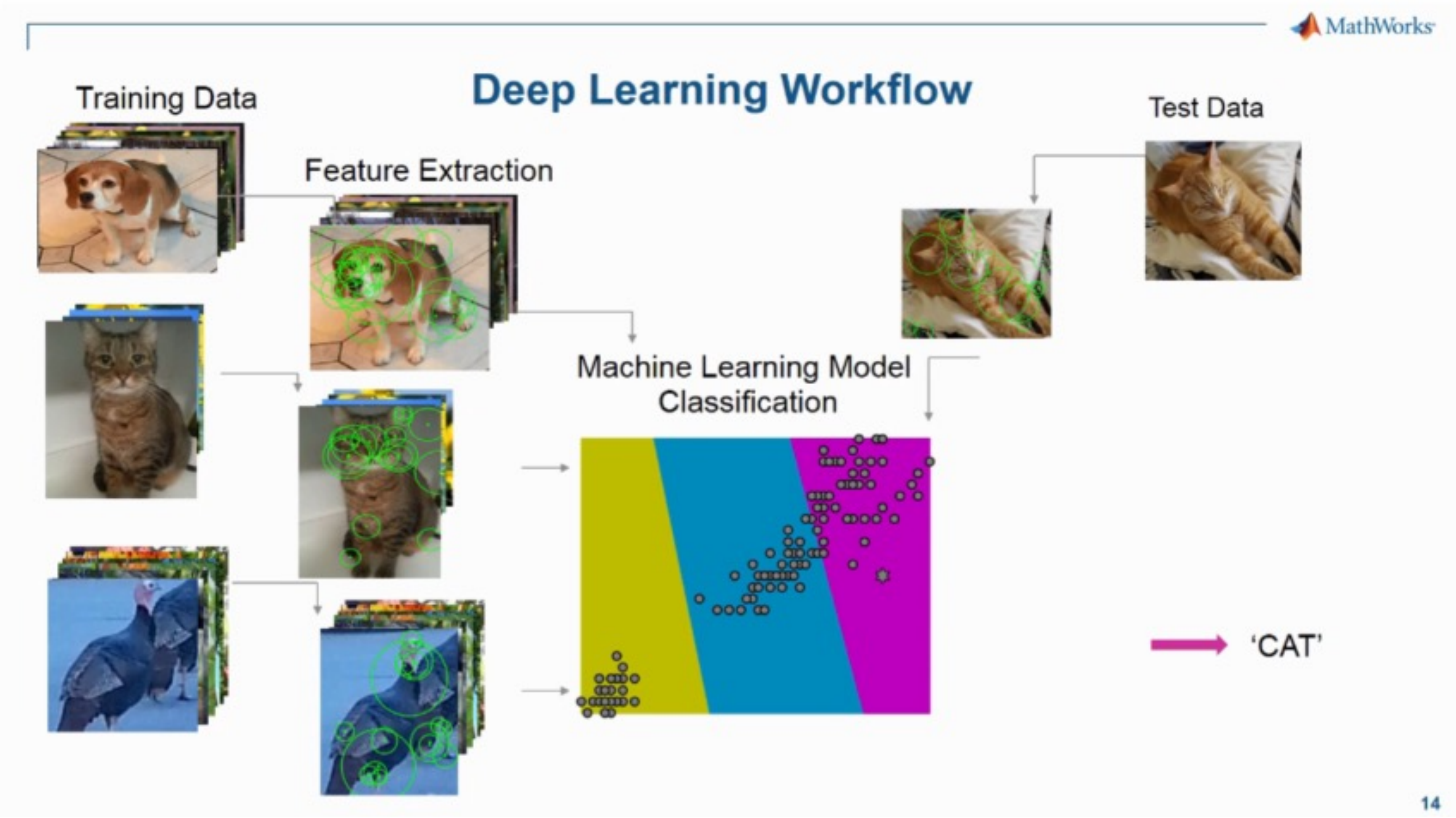
- Modeling the relationship between features of the data and some associated labels
- Trained model ----> usable to apply labels to new, unknown data
- **Classification** (discrete labels) and **regression** (continue labels)

Formally:

-Features set X; Label set Y

-Unknown target function f : X → Y

-Learning algorithm A uses the training set to select a function that approximate f (usually, there is a loss function that have to be minimized)



Deep Learning Workflow

Training Data

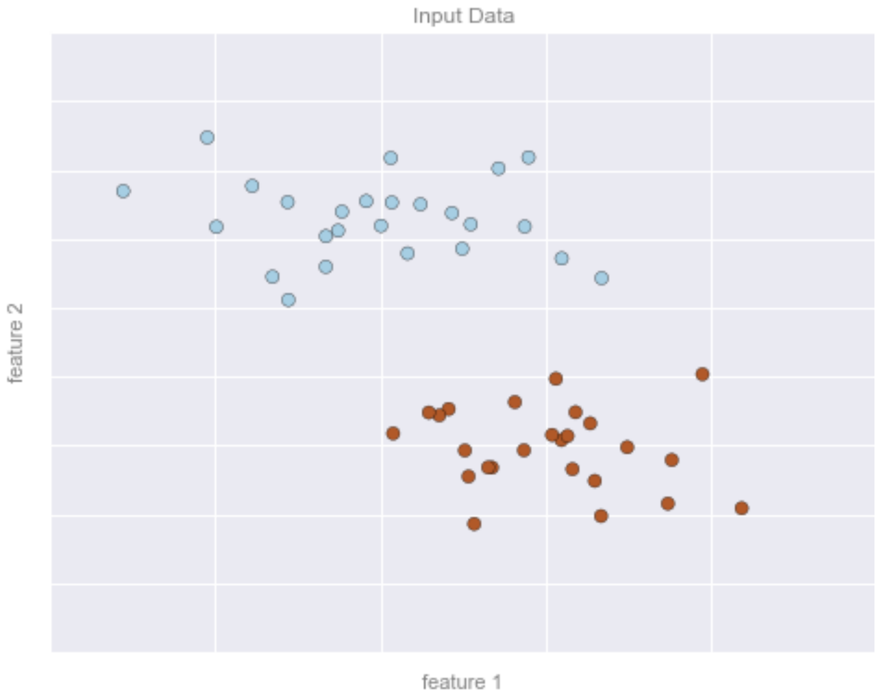
Feature Extraction

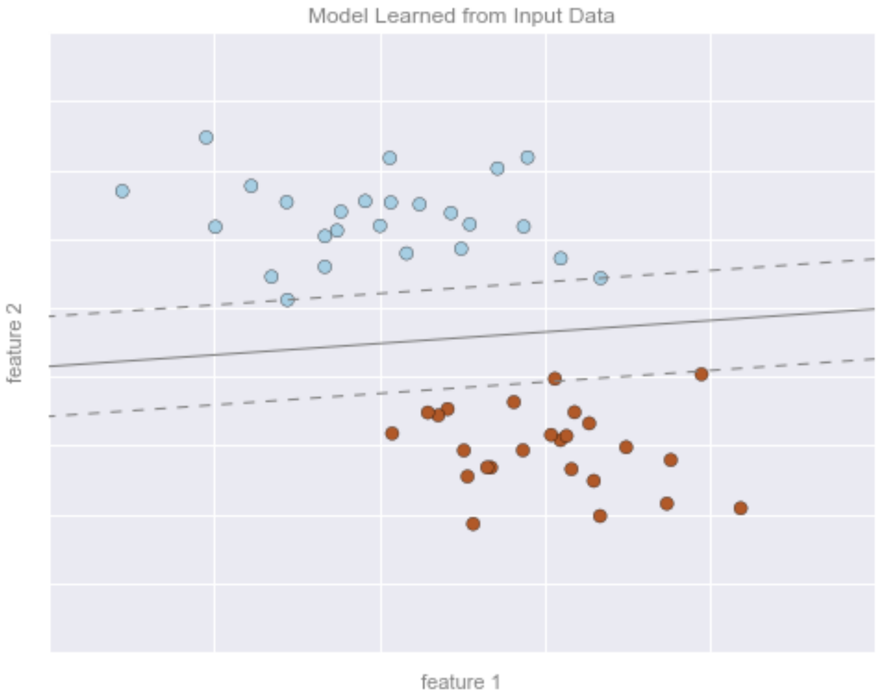
Machine Learning Model
Classification

Test Data

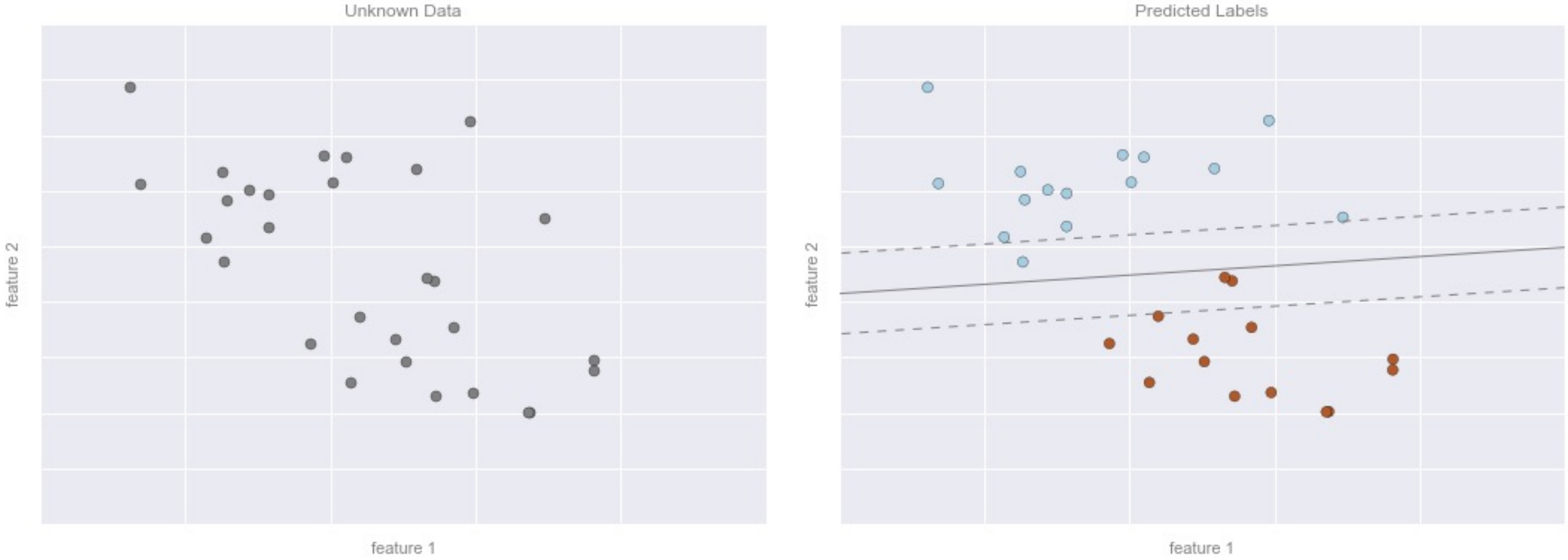
'CAT'

Classification

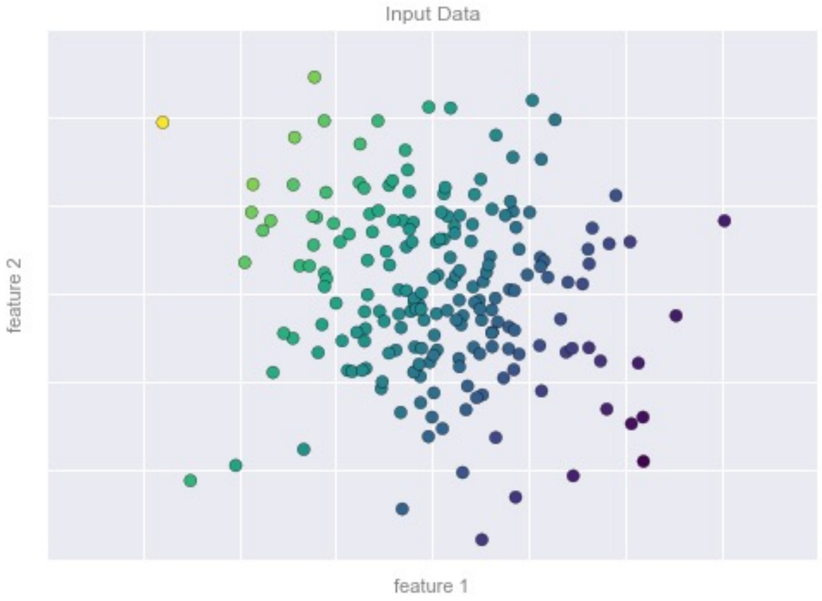


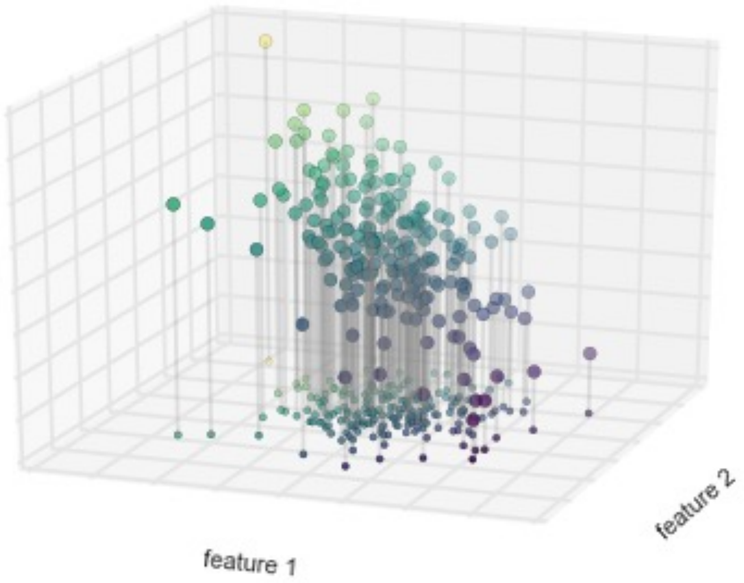


Classification

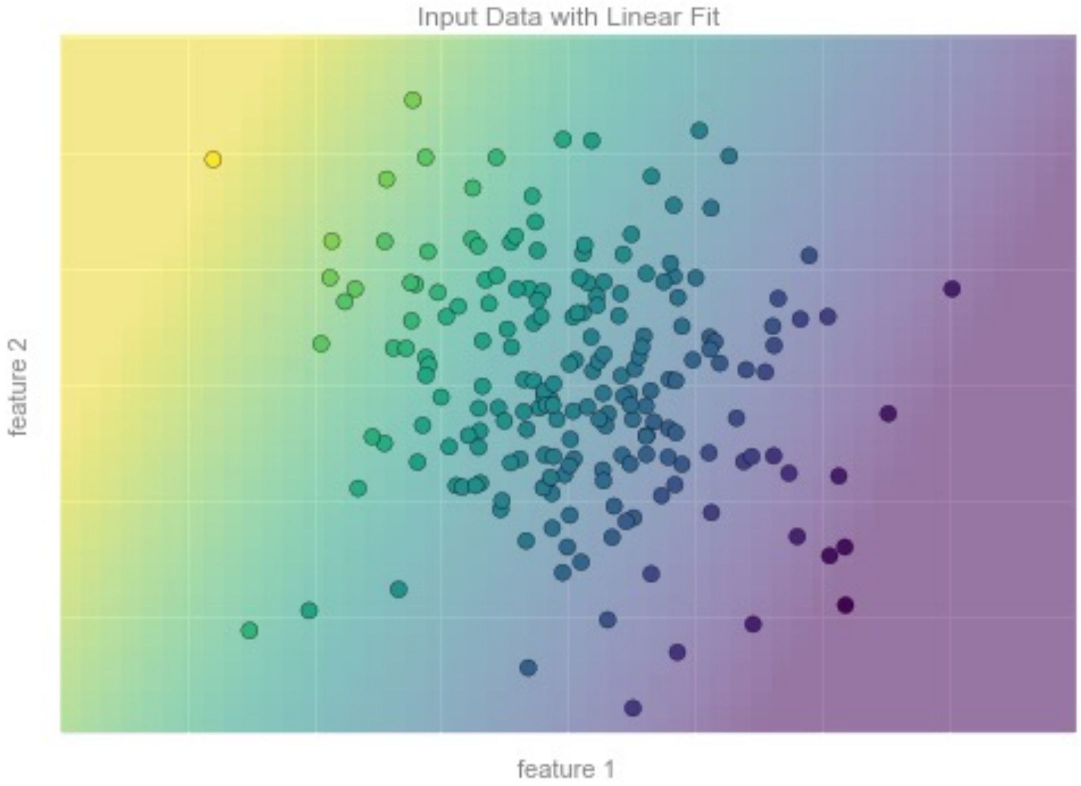


Regression

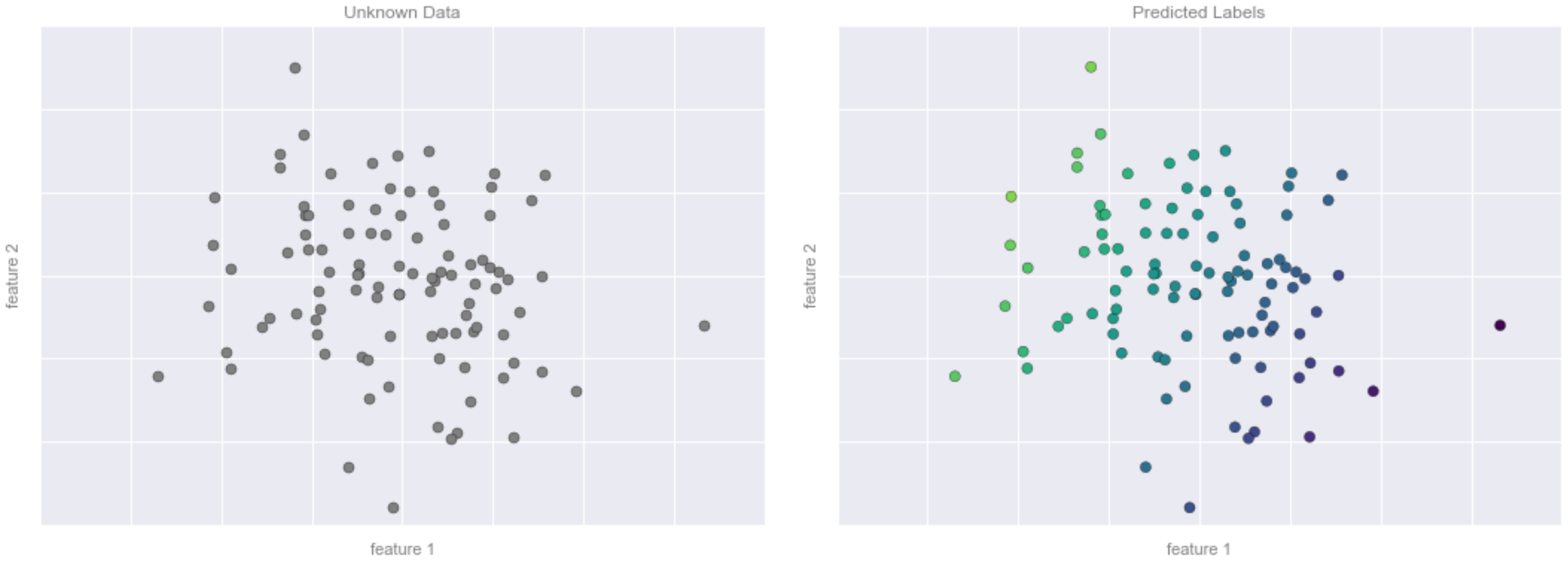




Regression



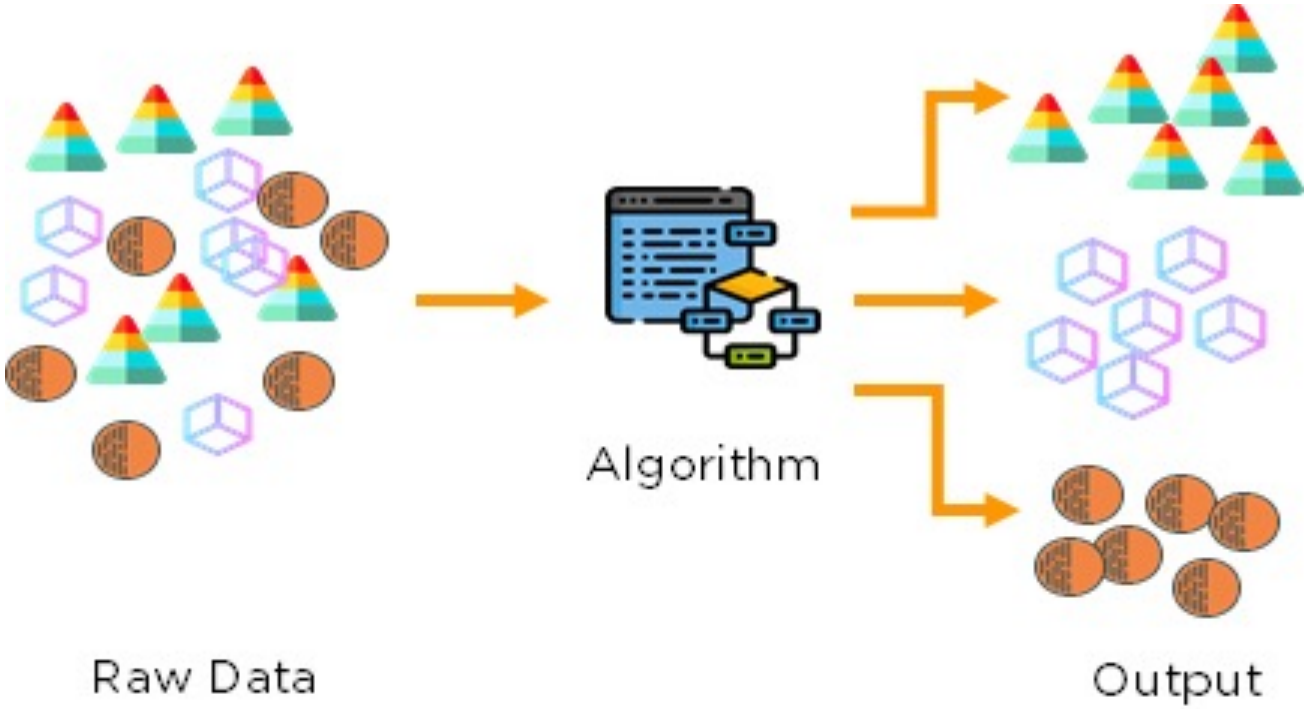
Regression


Unknown Data
feature 2
feature 1
Predicted Labels
feature 2
feature 1

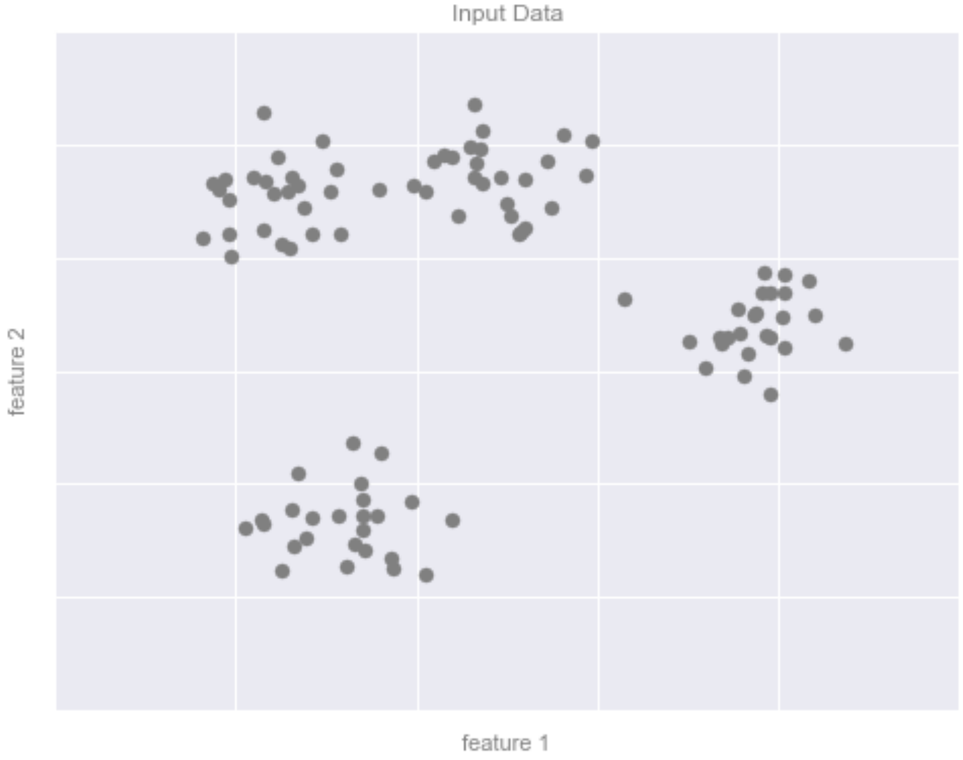
Unsupervised Machine Learning

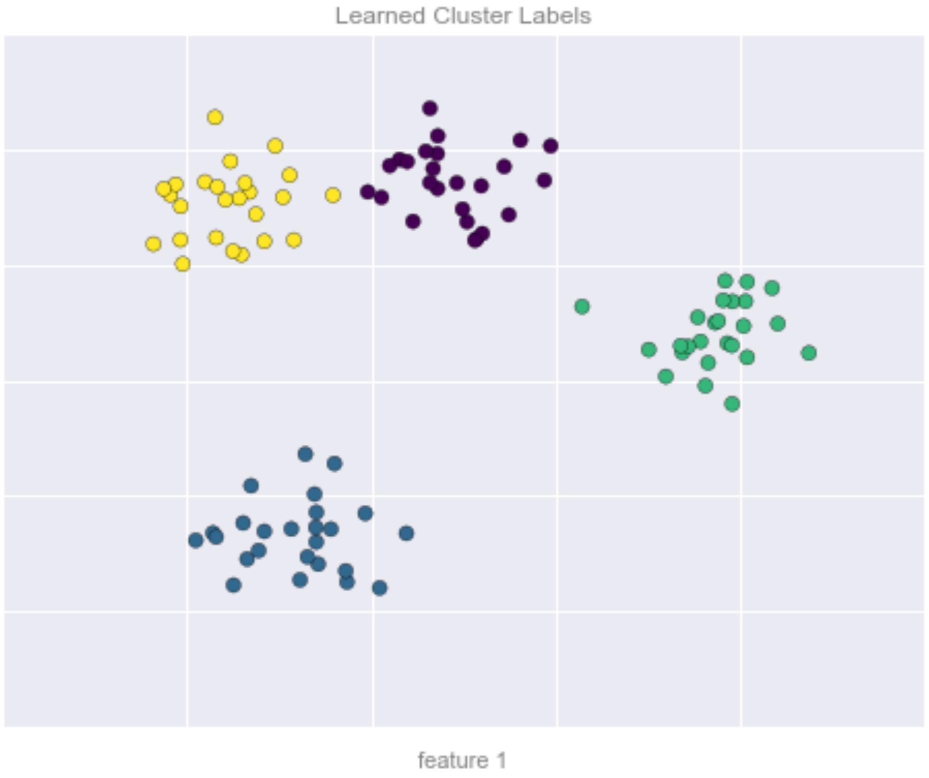
- Modeling the features of a dataset without reference to any label
- "Let the dataset speak for itself"

-Given an input Set X: find patterns, classify input into categories

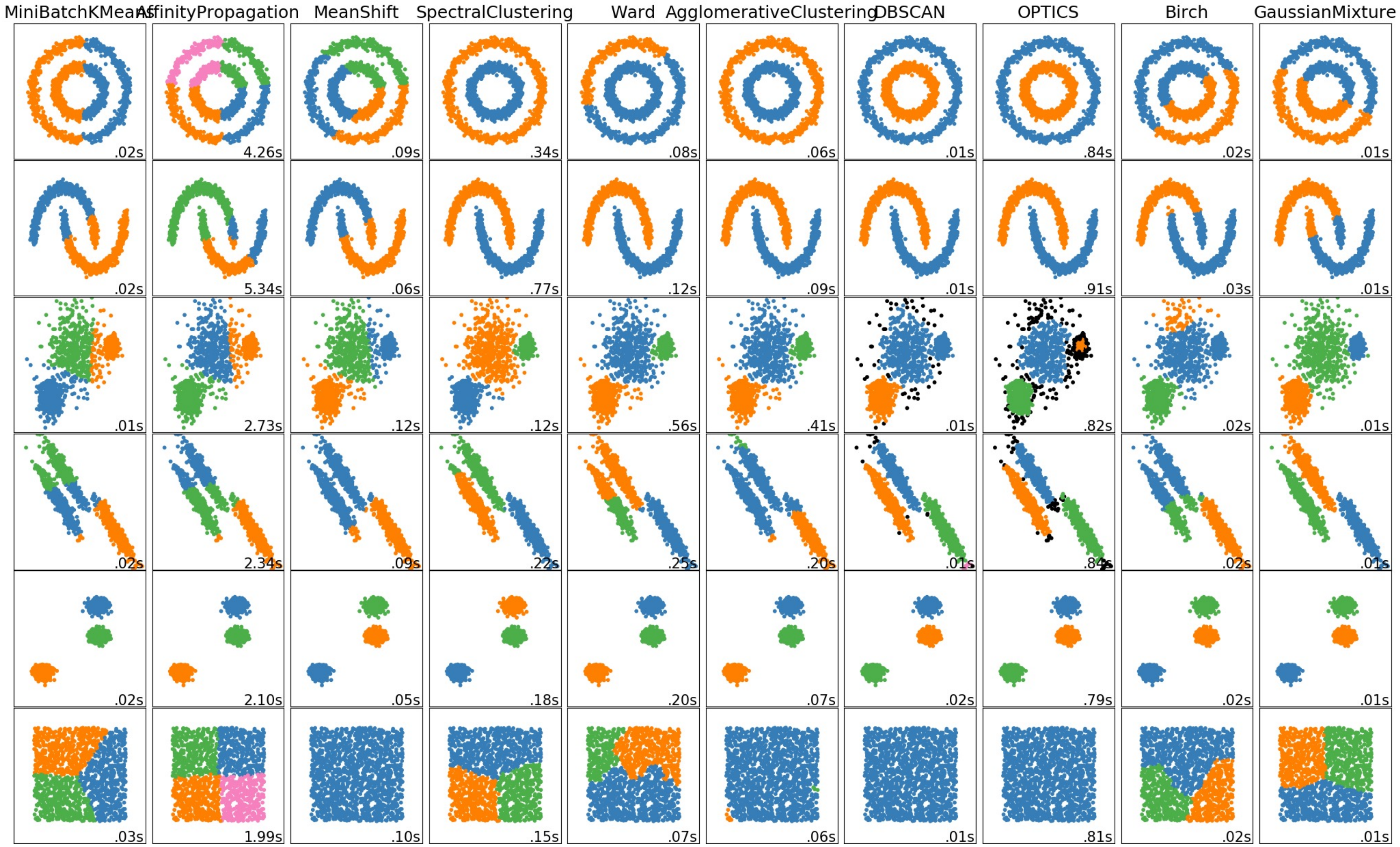

Raw Data
Algorithm
Output

Clustering


Input Data


Learned Cluster Labels

Clustering


MiniBatchKMeans
AffinityPropagation
MeanShift
SpectralClustering
Ward
AgglomerativeClustering
DBSCAN
OPTICS
Birch
GaussianMixture
.02s
4.26s
.09s
.34s
.08s
.06s
.01s
.84s
.02s
.01s
.02s
5.34s
.06s
.77s
.12s
.09s
.01s
.91s
.03s
.01s
.01s
2.73s
.12s
.12s
.56s
.41s
.01s
.82s
.02s
.01s
.02s
2.34s
.09s
.22s
.25s
.20s
.01s
.84s
.02s
.01s
.02s
2.10s
.05s
.18s
.20s
.07s
.02s
.79s
.02s
.01s
.03s
1.99s
.10s
.15s
.07s
.06s
.01s
.81s
.02s
.01s

Some Machine Learning Methods:

Supervised Machine Learning

-Logistic Regression

-Random Forest

-Support Vector Machine

-Naive Bayes

Unsupervised Machine Learning:

-K-Means (for clustering)

for Topic Modeling:

-LDA Algorithm

Logistic Regression

- The probabilities describing the possible outcomes of a single trial are modeled using a **logistic function**:

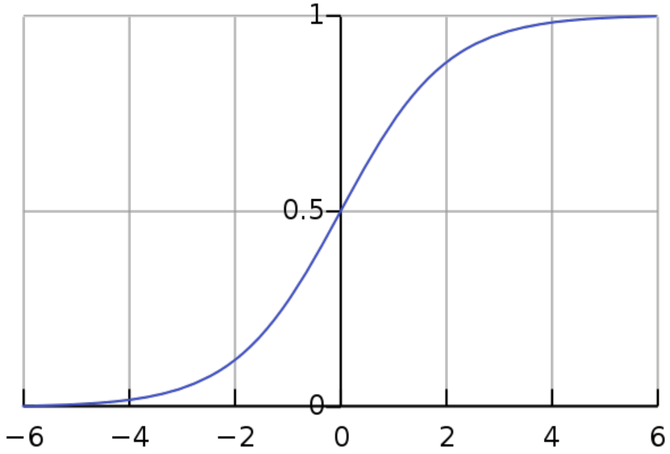
$$f(x) = \frac{L}{1 + e^{-k(x - x_0)}}$$

where x_0 is the x value of the sigmoid's midpoint

L is the curve's maximum value

k is the logistic growth rate or steepness of the curve

(The standard logistic function, with L=1, k=1 and x_0=0 is called the sigmoid)



Standard logistic function

Regularized logistic regression:

- Add a term in your minimization problem that give a "cost" to the number of coefficient that are different from zero (or give a cost the higher the size of the coefficients)

(Lasso, Ridge Regressions, etc)

When you implement logistic regression in scikit-learn, the default is to apply the L2 regularization (i.e. a Ridge regression)

Regularized logistic regression is used also to interpret the most predictive words for a category, calculating the **marginal effects** for each word and listing the most predictive words for each category.

Random Forest

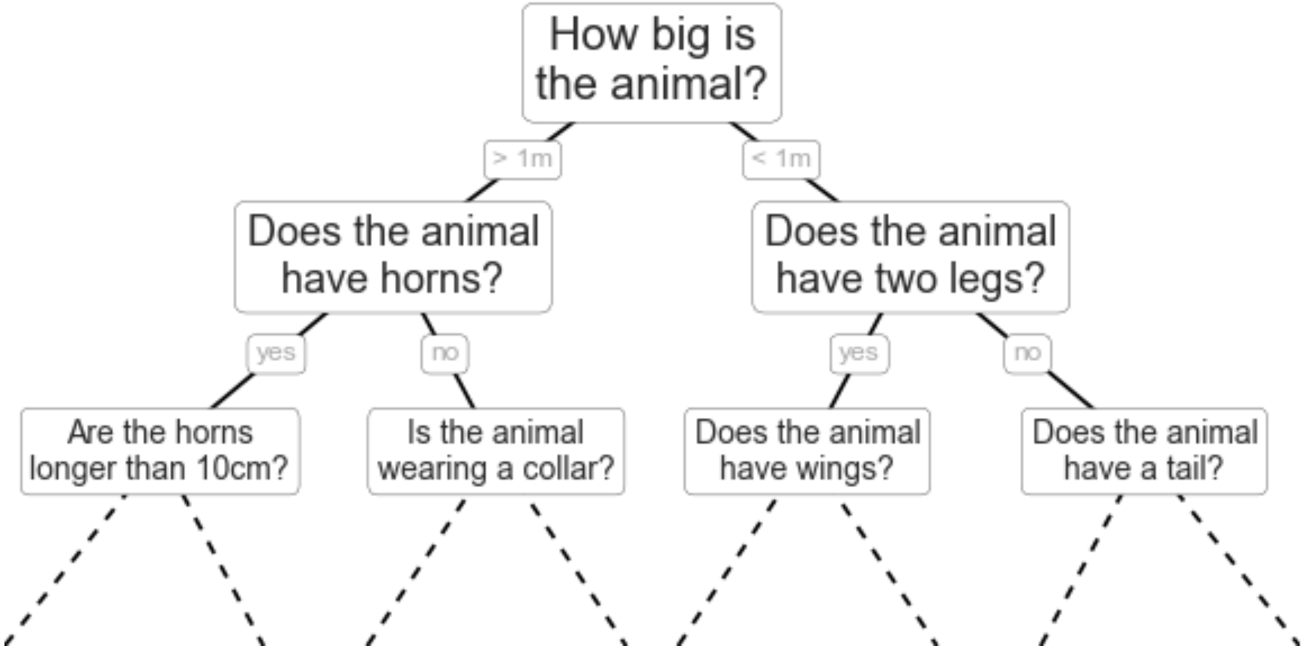
- Ensemble learner built on decision trees

Ensemble methods: they relies on aggregating the result of an ensemble of simpler estimators

What are Decision trees?

- Methods to classify data, based on "asking a series of questions" with a binary outcome
- Very intuitive (and interpretable) way to classify and label data

Decision Trees


How big is the animal?
> 1m
< 1m
Does the animal have horns?
Does the animal have two legs?
yes
no
yes
no
Are the horns longer than 10cm?
Is the animal wearing a collar?
Does the animal have wings?
Does the animal have a tail?

- Decision trees tend to overfit the data
- **Random forest:** ensable of decision trees (each tree in the ensemble is built from a sample drawn with replacement (bootstrap sample) from the training set

K-means

- Used for clustering
- Searches for a pre-determined number of clusters within an unlabeled multidimensional dataset

Optimal clustering:

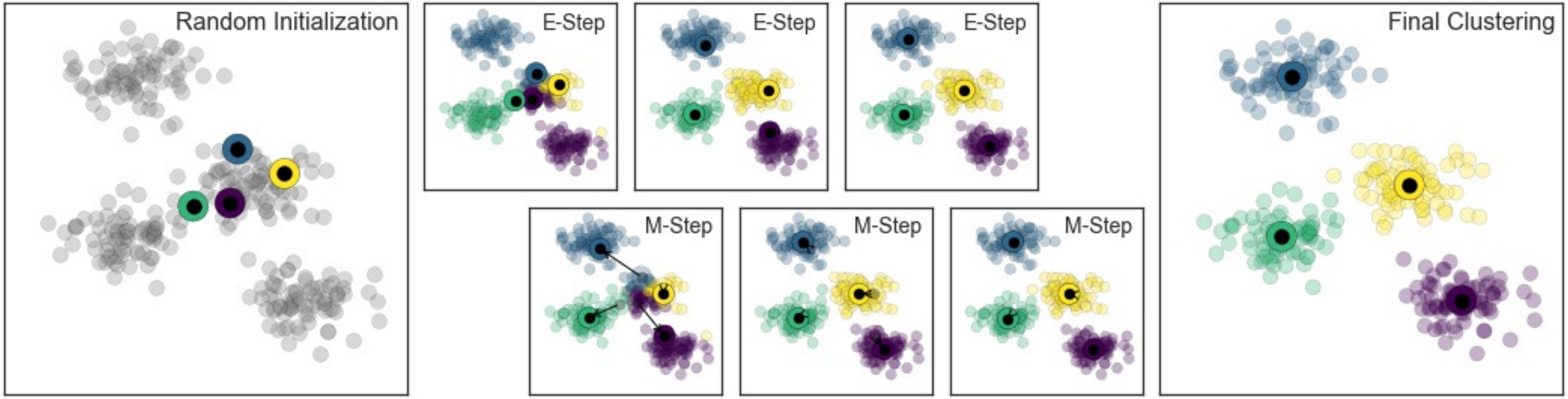
- The "cluster center" is the arithmetic mean of all the points belonging to the cluster
- Each point is closer to its own cluster center than to other cluster centers.

K-means

How does the algorithm work?

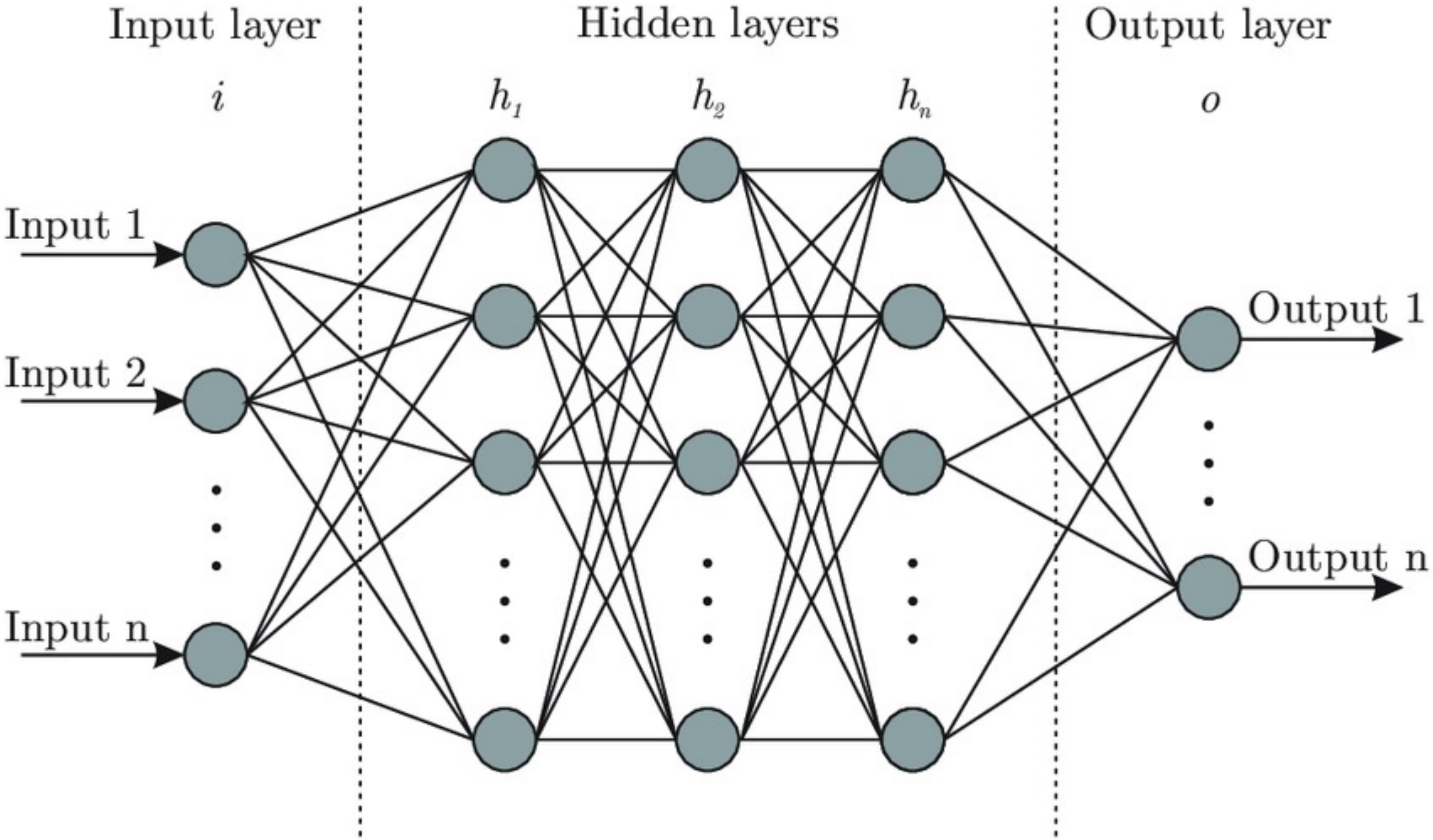
Expectation-maximization algorithm:

1) Guess some cluster centers
2) Repeat until converged
    a) E-Step: assign points to the nearest cluster center
    b) M-Step: set the cluster centers to the mean



A fast (very fast….) glimpse of Deep Learning

- Part of machine learning methods, based on artificial neural networks
- Inspired by how the brain works
- The frontier, for the majority of tasks in which machine learning is employed
- Applications: computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection, board game programs, …

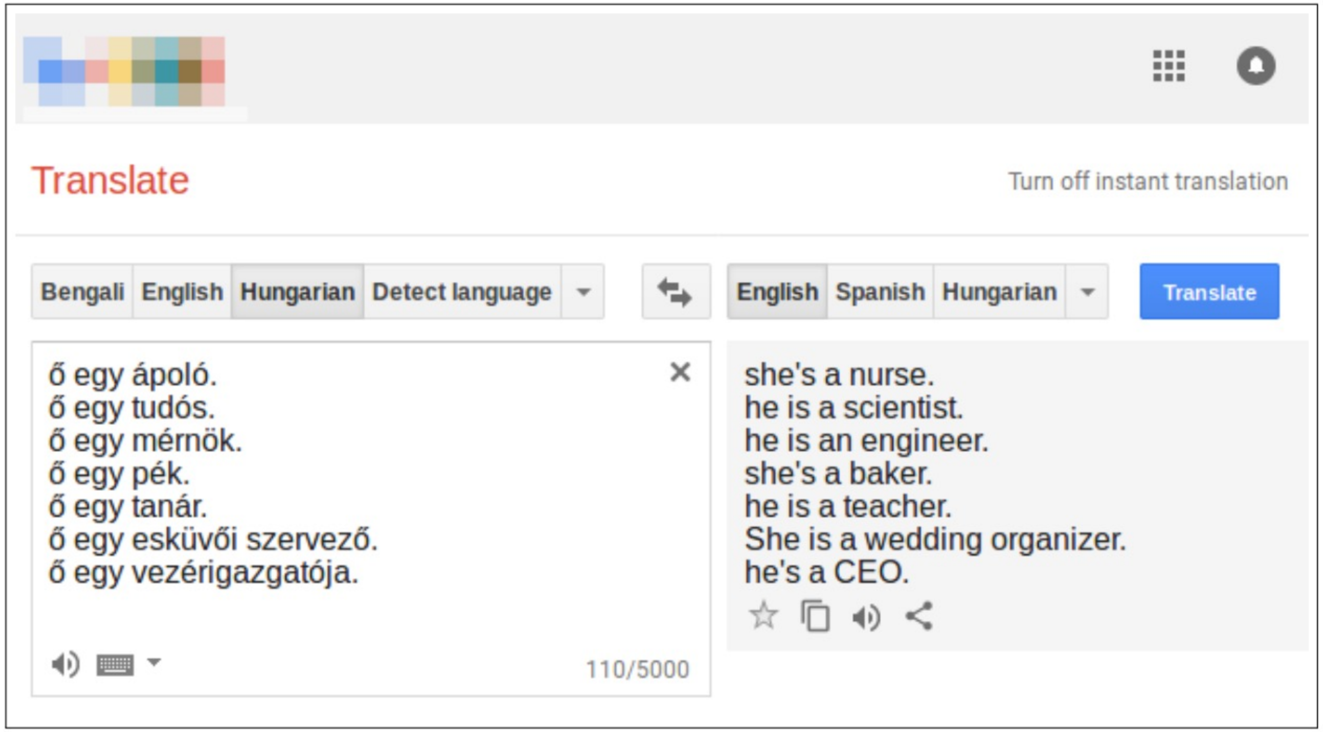

Input layer
Hidden layers
Output layer
i
h_1
h_2
h_n
o
Input 1
Input 2
Input n
Output 1
Output n

Problems of Machine Learning Algorithms:

-Interpretability

-They made persist bias that could exist in the dataset that is used for training

-Privacy



ő egy ápoló.
ő egy tudós.
ő egy mérnök.
ő egy pék.
ő egy tanár.
ő egy esküvői szervező.
ő egy vezérigazgatója.

she's a nurse.
he is a scientist.
he is an engineer.
she's a baker.
he is a teacher.
She is a wedding organizer.
he's a CEO.

Figure 1: Translating sentences from a gender neutral language such as Hungarian to English provides a glimpse into the phenomenon of gender bias in machine translation. This screenshot from Google Translate shows how occupations from traditionally male-dominated fields [40] such as scholar, engineer and CEO are interpreted as male, while occupations such as nurse, baker and wedding organizer are interpreted as female.

From the paper "Assessing Gender Bias in Machine Translation – A Case Study with Google Translate" (Prates, Avelar and Lamb, 2019). Link here: https://arxiv.org/pdf/1809.02208.pdf

How do economists use Machine Learning algorithms?

Mainly, **two ways** in which algorithms are used in economic analysis and research:

-To obtain (and use) econometrics methods that can give more information than usual econometrics models (i.e. random forest for estimation of heterogeneous treatment effects)

-To obtain information from "unstructured data" (i.e. videos, images, texts, etc) and use these information in their analysis (i.e. "Text-as-data" type of use)

Machine Learning for Econometrics

(I will just quote quickly this topic, but potentially, there is a growing literature on this topic, and some codes are already available both in R and in Python!!)

- **Theoretical/Conceptual papers**

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), 1228-1242.

Athey, S. (2015, August). Machine learning and causal inference for policy evaluation. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 5-6).

Lechner, M. (2018). Modified causal forests for estimating heterogeneous causal effects. *arXiv preprint arXiv:1812.09487*.

- **Applications**

Athey, S., & Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational Studies*, *5*(2), 37-51.

Goller, D., Harrer, T., Lechner, M., & Wolff, J. (2021). Active labour market policies for the long-term unemployed: New evidence from causal machine learning. *arXiv preprint arXiv:2106.10141*.

Cockx, B., Lechner, M., & Bollens, J. (2019). Priority to unemployed immigrants? A causal machine learning evaluation of training in Belgium. *arXiv preprint arXiv:1912.12864*.

Text-as-data Papers

(i.e. papers that use text to obtain some variables for their research question).

We are going to have a brief overview of the following papers:

- Ash, E., Chen, D. L., & Ornaghi, A. (2021**). Gender attitudes in the judiciary: Evidence from US circuit courts.** (R&R at *American Economic Journal: Applied Economics)*
- Cagé, J., Hervé, N., & Viaud, M. L. (2020**). The production of information in an online world: Is copy right?.** *The Review of Economic Studies, 87(5)*
- Sockin, J. (2021**). Show Me the Amenity: Are Higher-Paying Firms Better All Around?.** *Available at SSRN*.
- Hansen, S., Ramdas, T., Sadun, R., & Fuller, J. (2021). ***The demand for executive skills*** (No. w28959). National Bureau of Economic Research.

Gender attitudes in the judiciary: Evidence from US circuit courts (Ash, Chen and Ornaghi, 2021)

Research question: do gender attitudes influence interactions with female judges in US Circuit Courts?

Data: The authored opinions of 139 judges in U.S. (total corpus of over 14 million sentences), and outcomes: reversals of district court decisions, opinion assignment, and citations + decisions in gender-related cases.

Identification strategy: Differences-in-differences. Exploit the quasi-random assignment of judges to cases (in addition on conditioning on judges' characteristics)

Results: higher-slant judges vote more conservatively in gender-related cases. They interact differently with female colleagues: they are more likely to reverse lower-court decisions if the lower-court judge is a woman than a man, are less likely to assign opinions to female judges, and cite fewer female-authored opinions.

Gender attitudes in the judiciary: Evidence from US circuit courts (Ash, Chen and Ornaghi, 2021)

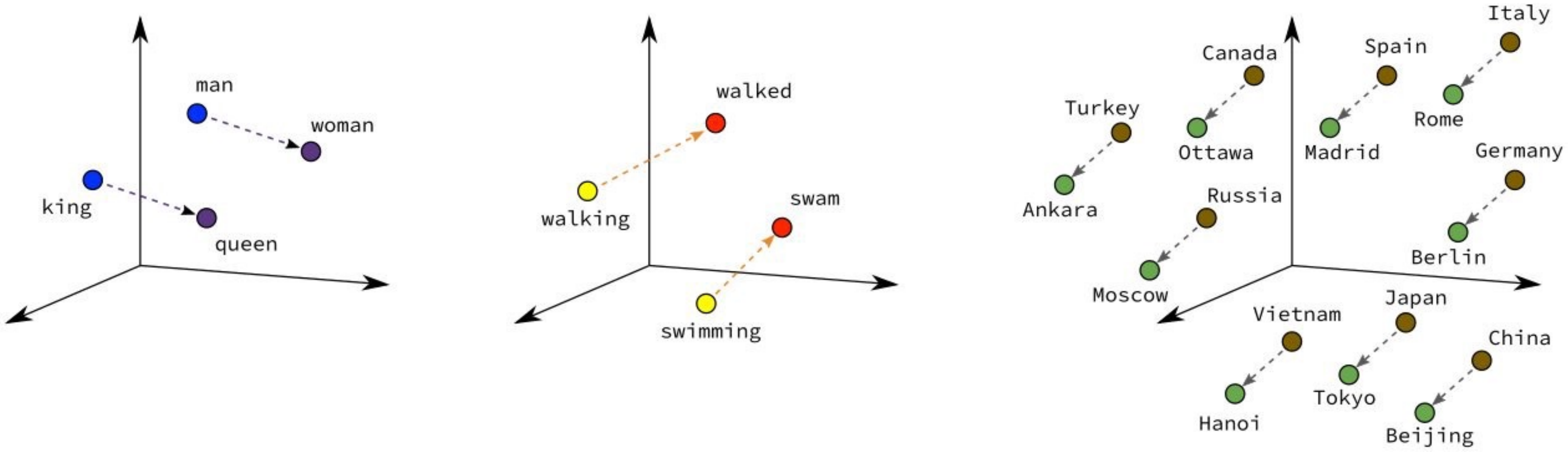
How does this paper exploit Machine Learning?

They construct a judge-specific measure of gender attitudes based on use of gender-stereotyped language in the judge's authored opinions.

----> They use NLP (natural language processing) to develop a measure of *gender slant* based on how strongly judges associate men with careers and women with families in the opinions they write

----> In particular: *word embedding* ----> an algorithm that distributes words in a vector space, based on their co-occurence in a corpus, and should represent their semantic meaning and relationships)

Word Embedding



Male-Female

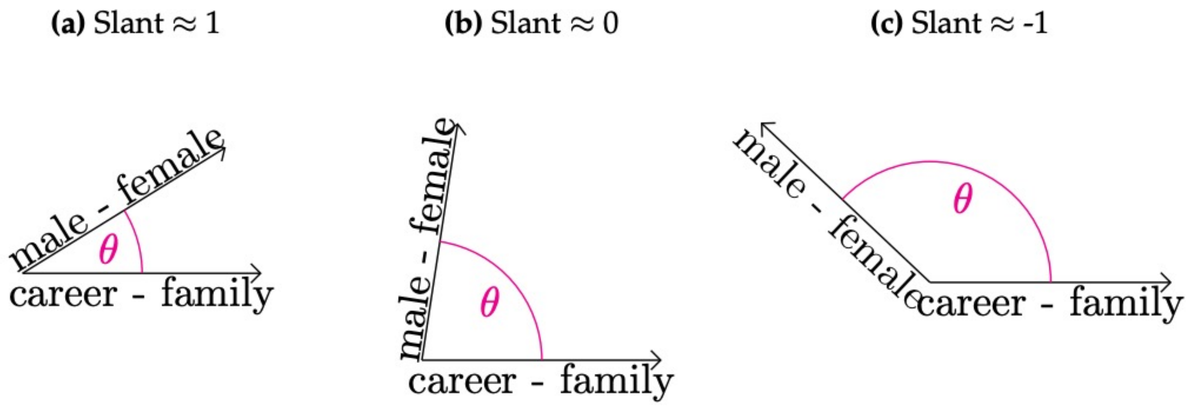
Verb Tense

Country-Capital

Gender attitudes in the judiciary: Evidence from US circuit courts (Ash, Chen and Ornaghi, 2021)

Cosine similarity ($\cos(\theta)$) between male – female (gender dimension) and career - family (career-family dimension)

Figure 2: Measuring Gender Attitudes using Cosine Similarity



(a) Slant $\approx$ 1

(b) Slant $\approx$ 0

(c) Slant $\approx$ -1

The Production of Information in an Online World" (Cagè, Hervè and Viaud, 2020)

Research Question: what it is the extent of copying in online media, and what are the estimations for the returns to originality in online news production?

Data: all the online content produced by French news media during 2013 + new micro audience data

Identification Strategy: (for return of originality): regression with event, date and media fixed effects

Results: they document high reactivity of online media: one quarter of the news stories are reproduced online in under 4 minutes. We show that this is accompanied by substantial copying, both at the extensive and at the intensive margins. Returns to originality: they find that original content producers tend to receive more viewers, thereby mitigating the newsgathering incentive problem raised by copying.

The Production of Information in an Online World" (Cagè, Hervè and Viaud, 2020)

How do they use Machine Learning?

They perform a **topic detection algorithm** to construct the set of news stories. Each document is placed within the most appropriate cluster, i.e. the one that discusses the same event-based story. They obtain a total number of 25,000 stories.

(+ document the propagation of the story, + plagiarism algorithm, + copying with acknowledgment)

Show Me the Amenity: Are Higher-Paying Firms Better All Around?
(Jason Sockin, 2021)

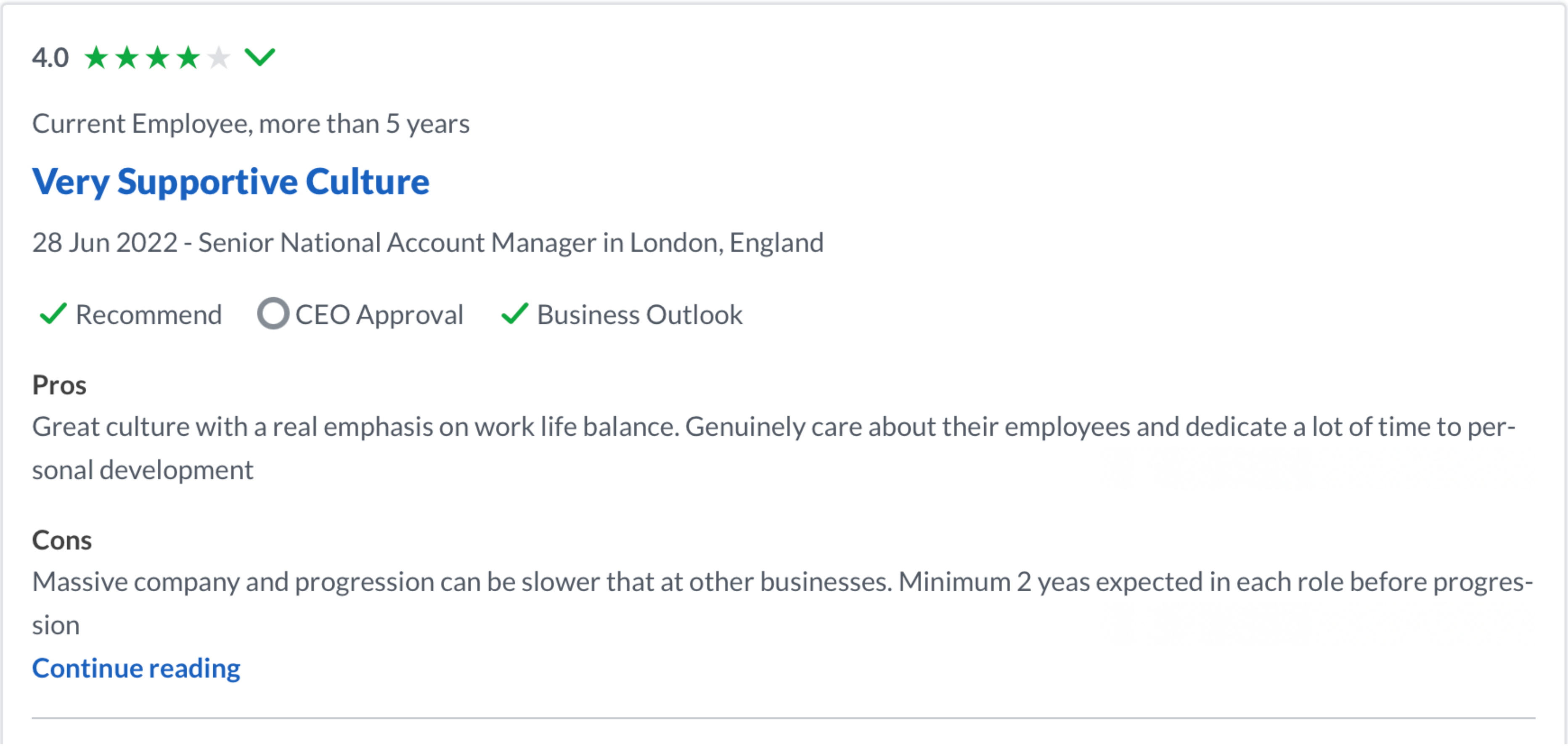
Research Question: Do higher-paying firms offer more favorable work, or compensate for less favorable work? (estimation of joint distribution of wages, amenities, and job satisfaction across firms)

Data: job reviews/reports from Glassdoors

Identification strategy: AKM (and other analysis)

Results: 1) High-paying firms are high-satisfaction firms because they offer better amenities 2) workers, especially high-earners, are willing to pay for job satisfaction 3) incorporating non-wage amenities nearly doubles the variance in total compensation across firms

Glassdoor review



4.0 ★★★★☆

Current Employee, more than 5 years

Very Supportive Culture

28 Jun 2022 - Senior National Account Manager in London, England

✓ Recommend ○ CEO Approval ✓ Business Outlook

Pros

Great culture with a real emphasis on work life balance. Genuinely care about their employees and dedicate a lot of time to personal development

Cons

Massive company and progression can be slower that at other businesses. Minimum 2 yeas expected in each role before progression

Continue reading

Show Me the Amenity: Are Higher-Paying Firms Better All Around? (Jason Sockin, 2021)

How do they use Machine Learning?

Semi-surpervised Topic Modeling (Anchored Correlation Explanation – CorEx model – of Gallagher et al. (2017) on workers' free response description (pros and cons) in Glassdoor

1) The researcher chose 48 amenities (that he thinks are important in these descriptions)
2) He gives to the algorithm some topic-specific "anchor words", to help guide topic to convergence (and improve interpretability of topics) (the anchor words still refer to the 48 amenities he identifies)
3) The algorithm create the clusters (according to its objective)

The Demand for Executive Skills (Hansen, Ramdas, Sadun and Fuller, 2021)

Research Question: which skills are required in managerial labour markets?

Data: large corpus of detailed and previously unexplored job descriptions for C-suite positions spanning a time period of 17 years (provided to them by one of the world's largest headhunting companies)

Results: the data show an increasing relevance of social skills in top managerial occupations, and a greater emphasis on social skills in larger and more information intensive organizations.

The Demand for Executive Skills (Hansen, Ramdas, Sadun and Fuller, 2021)

How do they use Machine Learning?

They classify the information contained in these documents using methods borrowed from machine learning, which allow us to map unstructured, free-text data into distinct clusters of skill requirements. They use the data to examine the variation in the demand for different managerial skills which provide the first direct evidence on C-suite skill requirements.

The Demand for Executive Skills (Hansen, Ramdas, Sadun and Fuller, 2021)

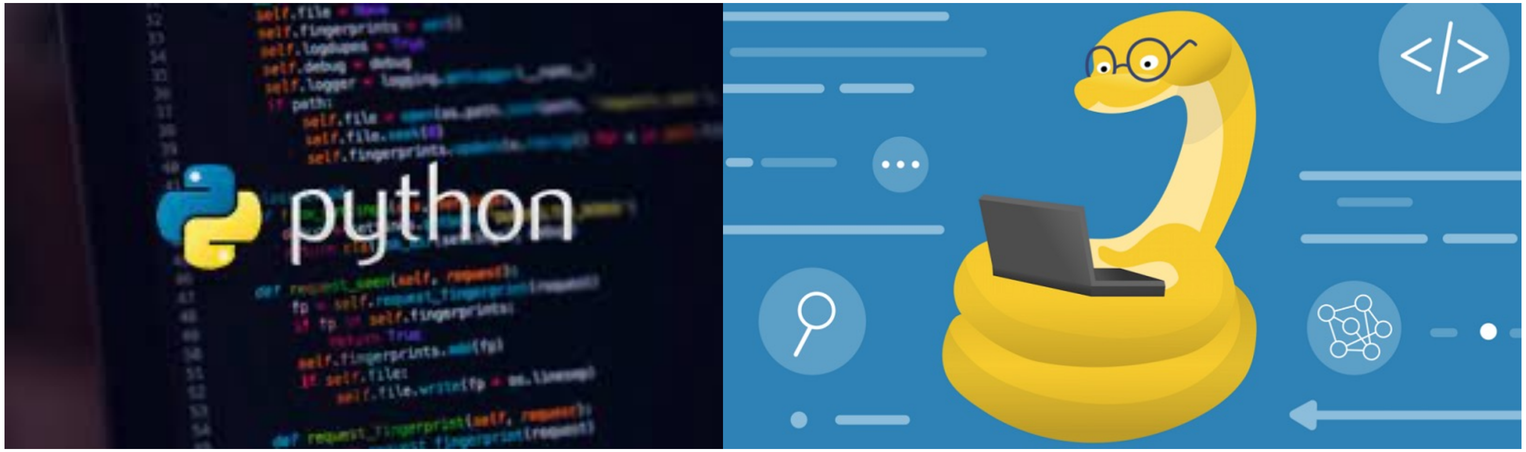
"We propose a novel classification approach to derive economically interpretable measures from the unstructured text of this corpus. Our approach involves two steps. First, we define a comprehensive vector of skills requirements that are relevant for Chief Executives. We obtain this by collecting the numerous textual descriptions of skills from the O*NET entry for the Chief Executive occupation, and clustering them into six broad categories using a k-means algorithm. Second, we express each job description in the search corpus in terms of the relative demand for each skill cluster by comparing the similarity of the language included in the document with the text associated with each of the O*NET clusters.

Both the clustering of O*NET skills into groups and the comparison of job texts to O*NET texts require the quantification of linguistic relatedness. We compute this via a language embedding model estimated from an auxiliary corpus of all Harvard Business Review articles from its inception in 1922 to the present day. This large, domain-specific corpus allows us to obtain semantic relationships between words in the context of business and management. We then apply the model to measure similarity in the O*NET and job search corpora, an approach known as transfer learning."

Getting our hands dirty: using Python as programming language

Why Python?

- High-level, interpreted, general-purpose programming language.
- Higher code readability (with respect to other coding languages)
- Great libraries for scientific computing and data science (NumPy, Pandas, Matplotlib), for machine learning (TensorFlow, Keras, Pytorch, Scikit-Learn) and for natural languange processing (NLTK, Gensim, SpaCy, etc)
- Some pretrained models available for download (to be fined-tuned) from platforms like HuggingFaces



Implementation of a Machine Learning pipeline: Python and the Scikit-Learn package

- Python is one of the most used programming language for easily implementing a Machine Learning pipeline

-----> Good packages for deep learning methods (TensorFlow, Keras), and some (large) pretrained models are made available online (i.e. on HuggingFaces)

- Scikit-Learn provides a "standard grammar" to implement models
- The pipelines to train algorithms are basically the same (to change model to be train is enough usually to change few or only a line)

Pipeline for implementing a Supervised Machine Learning Algorithm:

-preprocessing (i.e. putting everything lowercases, stemming, removing stopwords, etc)

-divide your dataset in traning and test dataset

On the training dataset:

-creating the **matrix of token counts**

(has as columns all the words of our vocabulary and as rows the collection of texts we are analyzing, and each element in each row represents the frequency of the word in the column in that text (*CounterVectorizer* command in Python).

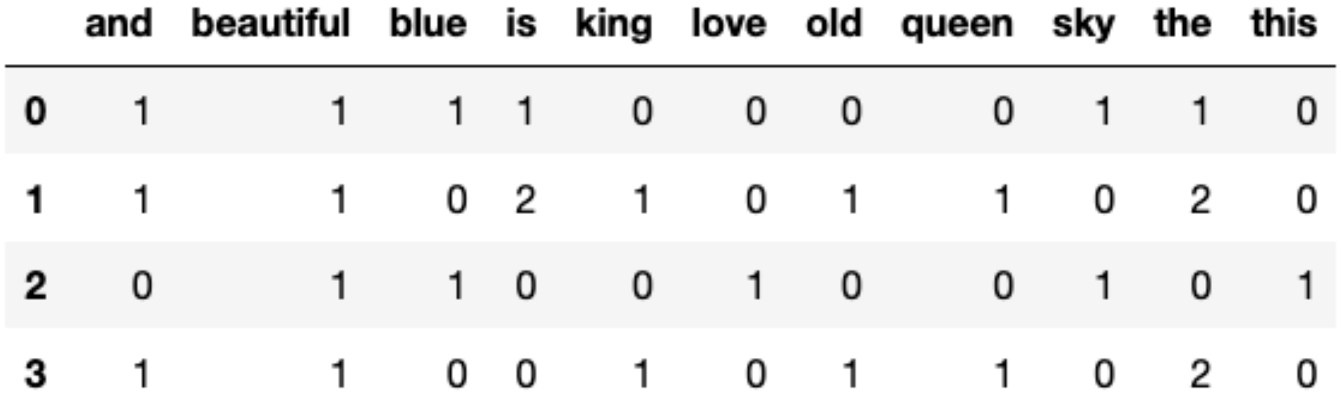
-Applying the **Tfidf transformation** (see next slides)

Preprocessing

(here, I removed some stopwords, I put everything in lowercases and I stemmed words).

Original text	**Pre-processed text**
Kinda sounds like a RED WAVE to me, huh? Which is exactly what I predicted. I went against every lying fake news poll. They were all wrong. I was right.	kinda sound like red wave me huh exactli predict went everi ly fake new poll wrong right
Republicans Are Joining CNN’s Ana Navarro For Meeting With Soros-Backed Amnesty Group	republican join cnn’ ana navarro meet sorosback amnesti group
Too funny James Woods	funni jame wood

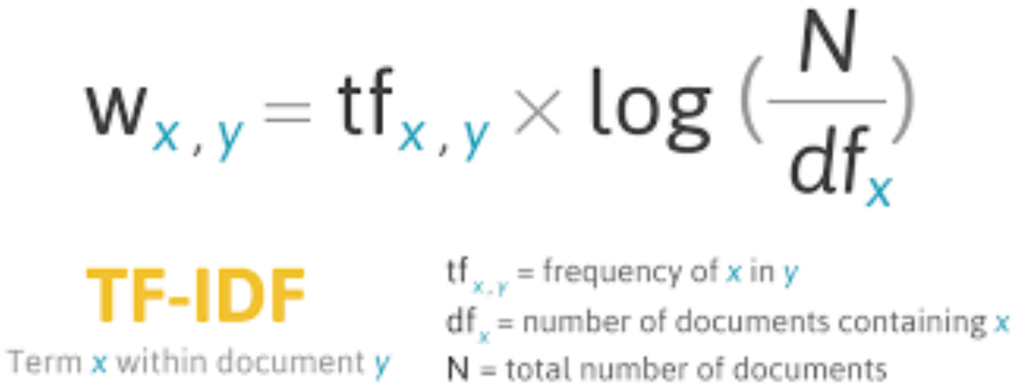
CounterVectorizer:



	and	beautiful	blue	is	king	love	old	queen	sky	the	this
0	1	1	1	1	0	0	0	0	1	1	0
1	1	1	0	2	1	0	1	1	0	2	0
2	0	1	1	0	0	1	0	0	1	0	1
3	1	1	0	0	1	0	1	1	0	2	0

Tfidf (term frequency-inverse document frequency)

- Weighting factor -----> reflect **how** important a word is to a document in a collection or corpus
- Increases proportionally to the number of times a word appears in the document
- Offset by the number of documents in the corpus that contain the word



$$w_{x,y} = tf_{x,y} \times \log\left(\frac{N}{df_x}\right)$$

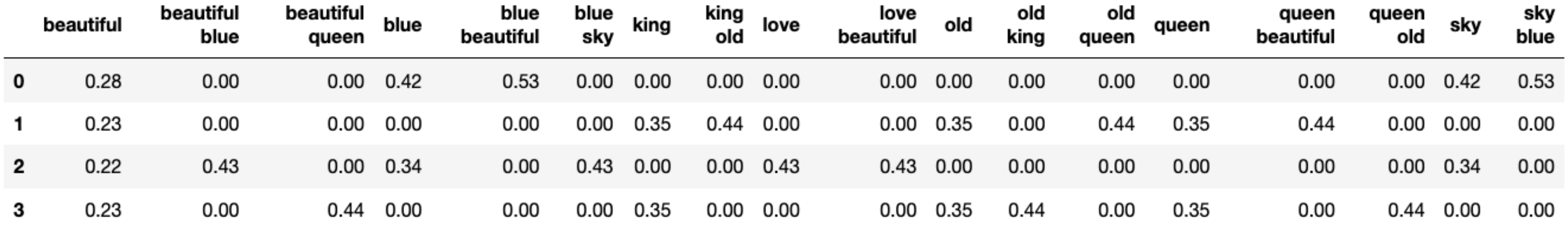
TF-IDF

Term x within document y

$tf_{x,y}$ = frequency of x in y

df_x = number of documents containing x

N = total number of documents



	beautiful	beautiful blue	beautiful queen	blue	blue beautiful	blue sky	king	king old	love	love beautiful	old	old king	old queen	queen	queen beautiful	queen old	sky	sky blue
0	0.28	0.00	0.00	0.42	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.53
1	0.23	0.00	0.00	0.00	0.00	0.00	0.35	0.44	0.00	0.00	0.35	0.00	0.44	0.35	0.44	0.00	0.00	0.00
2	0.22	0.43	0.00	0.34	0.00	0.43	0.00	0.00	0.43	0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.34	0.00
3	0.23	0.00	0.44	0.00	0.00	0.00	0.35	0.00	0.00	0.00	0.35	0.44	0.00	0.35	0.00	0.44	0.00	0.00

(Continue) Pipeline for Supervised Machine Learning Learning:

- Select your model
- Fit your model on your training dataset
- Prepare your test set to be readable from the algorithm (i.e. apply the CounterVectorizer fitted on the data of the training dataset, and then apply tfidf on the matrix obtained)
- Predict your labels applying the algorithm on the test set
- Cross-validation
- Calculate metrics of algorithms performance based on your results on the test set
- Try other models, and see which one perform better

Pipeline for Unsupervised Machine Learning Methods:

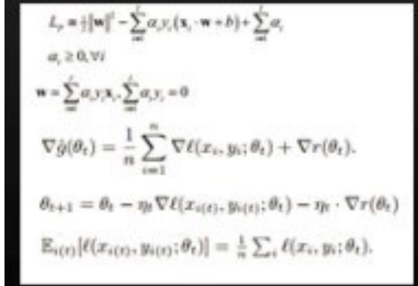
- Preprocessing
- CounterVectorized and Tfidf Transformation
- Choose your model (and the associated parameters)
- Apply your model to your data
- Analyze ("manually") the groups you obtained

Machine Learning

what society thinks I do

what my friends think I do

what my parents think I do



$$L_p = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l}\alpha_i y_i(\mathbf{x}_i \cdot \mathbf{w} + b) + \sum_{i=1}^{l}\alpha_i$$

$$\alpha_i \geq 0, \forall i$$

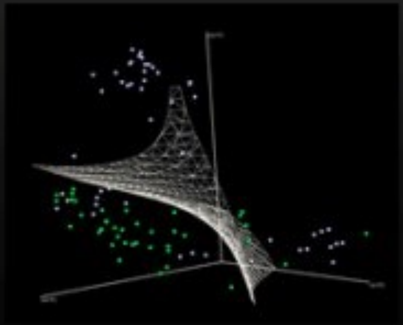
$$\mathbf{w} = \sum_{i=1}^{l}\alpha_i y_i \mathbf{x}_i, \sum_{i=1}^{l}\alpha_i y_i = 0$$

$$\nabla\hat{g}(\theta_t) = \frac{1}{n}\sum_{i=1}^{n}\nabla\ell(x_i, y_i; \theta_t) + \nabla r(\theta_t).$$

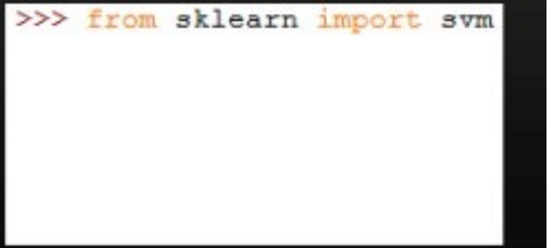
$$\theta_{t+1} = \theta_t - \eta_t \nabla\ell(x_{i(t)}, y_{i(t)}; \theta_t) - \eta_t \cdot \nabla r(\theta_t)$$

$$\mathbb{E}_{i(t)}[\ell(x_{i(t)}, y_{i(t)}; \theta_t)] = \frac{1}{n}\sum_i \ell(x_i, y_i; \theta_t).$$

what other programmers think I do



what I think I do



```
>>> from sklearn import svm
```

what I really do

Let's now go to the notebook to see concretely how these steps are implemented….. :)

References

Ash, E., Chen, D. L., & Ornaghi, A. (2021**). Gender attitudes in the judiciary: Evidence from US circuit courts.** (R&R at *American Economic Journal: Applied Economics)*

Cagé, J., Hervé, N., & Viaud, M. L. (2020**). The production of information in an online world: Is copy right?.** *The Review of Economic Studies, 87(5)*

Hansen, S., Ramdas, T., Sadun, R., & Fuller, J. (2021). ***The demand for executive skills*** (No. w28959). National Bureau of Economic Research.

Python Data Science Handbook (by Jake VanderPlas, https://jakevdp.github.io/PythonDataScienceHandbook/)

Sockin, J. (2021**). Show Me the Amenity: Are Higher-Paying Firms Better All Around?.** *Available at SSRN*.