

CSPB 3202 Artificial Intelligence

2020 Spring

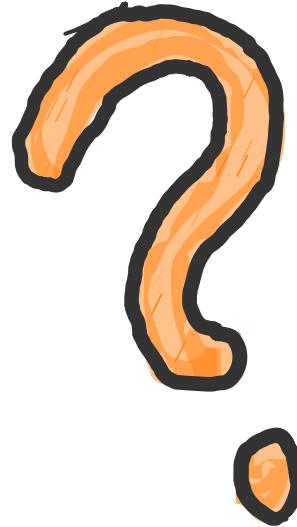
Machine Learning

Geena Kim



University of Colorado Boulder

The BUZZ words



Data SCience

Artificial Intelligence

Machine Learning

Deep Learning

Data Science

- Interdisciplinary field concerning Data
- Almost anything to do with Data:
(e.g.) Data Pipelining, Data Munging, Data Analysis, EDA...
- Includes Soft or Hard science, Small or Big data
- On the Job, Data Scientists do...

What does a Data Scientist do?

Data Scientists are responsible for the collection, cleaning and munging of data to meet the company's purpose. Duties vary according to the industry and may include experimental frameworks for product development and machine learning with the aim to lay a strong data foundation for robust analytics to be performed.

source: indeed.com

Artificial Intelligence

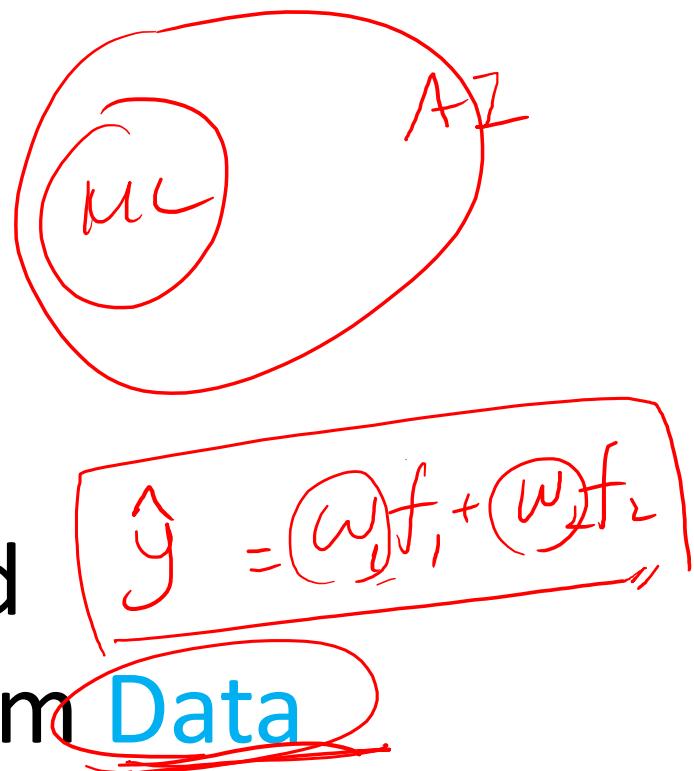
- One of the oldest and core subjects in CS
- About Problem-solving with Intelligence
- Theoretical and Practical
- On the Job, AI engineers/experts need skills..

Math and programming skills, Problem-solving skills, Machine learning....

and work on...

Build and deploy an AI system, build ML models, NLP, Robotics, Computer Vision...

Machine Learning



- Subfield of Artificial Intelligence
- About various statistical models and learning algorithms and training from Data
- Supervised/Unsupervised Learning / RL
- Modern version of Statistical Learning
- On the Job, ML engineers do..

Develop and test ML models, Design ML experiments, build ML system....

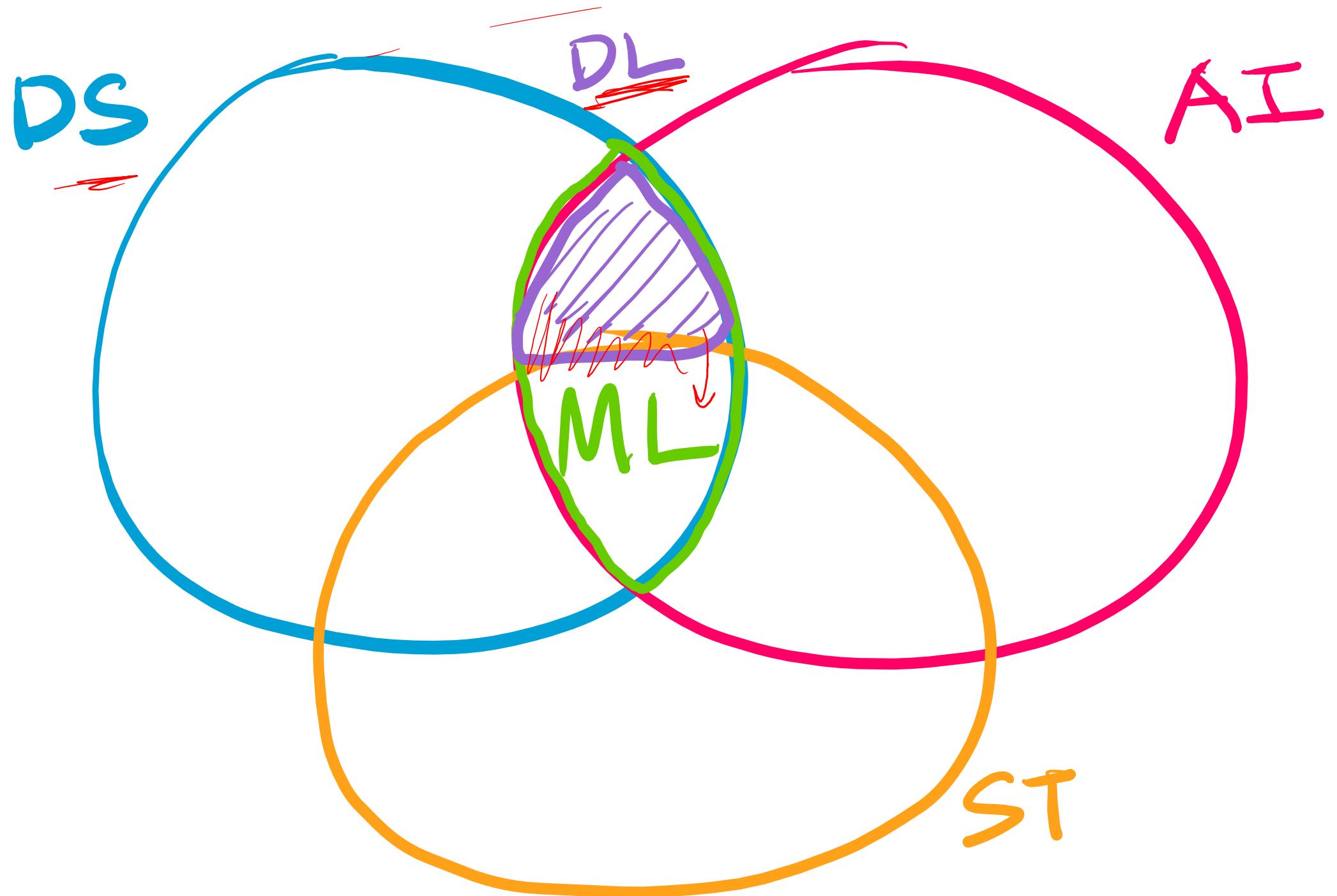
Deep Learning



- About Neural Network models, model training from Data, and all kinds of training techniques
- Subfield of Artificial Intelligence
- Subfield of Machine Learning
- On the Job, DL engineers do..

Solving complex technical challenges in various areas of deep learning such as object detection, segmentation, video understanding, sequence prediction, adaptive computing, memory networks, reduced precision training and inference, graph compilers, reinforcement learning, search distributed and federated training, and more

source: NVIDIA



Good Time to Learn Machine Learning!

Machine learning
Field of study

Software engineering
Field of study

+ Add comparison

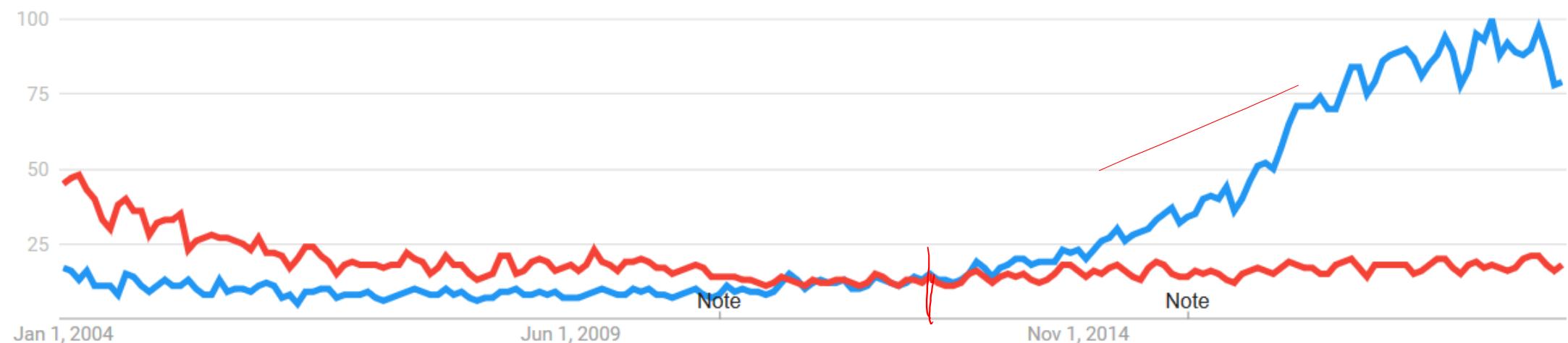
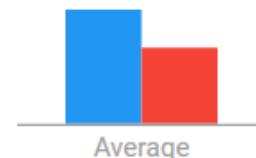
United States ▾

2004 - present ▾

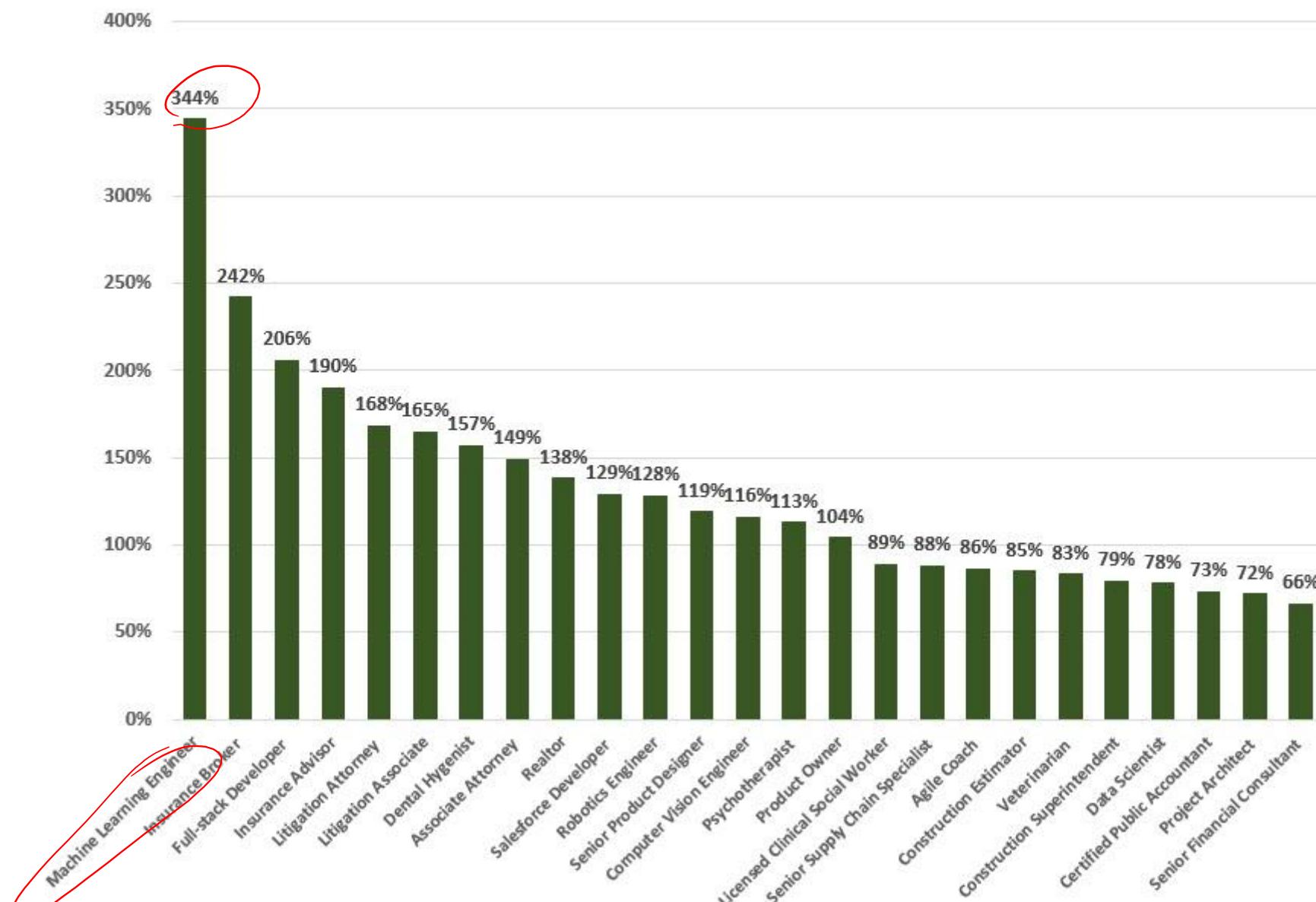
All categories ▾

Web Search ▾

Interest over time ?



Indeed's Best Jobs In The U.S.
% Growth in # of postings, 2015 - 2018
March 14, 2019



Ok, ML sounds Cool !

What Can I do with ML ?

Machine Learning is Everywhere



Unsupervised

Product Recommendations



Inspired by your shopping trends

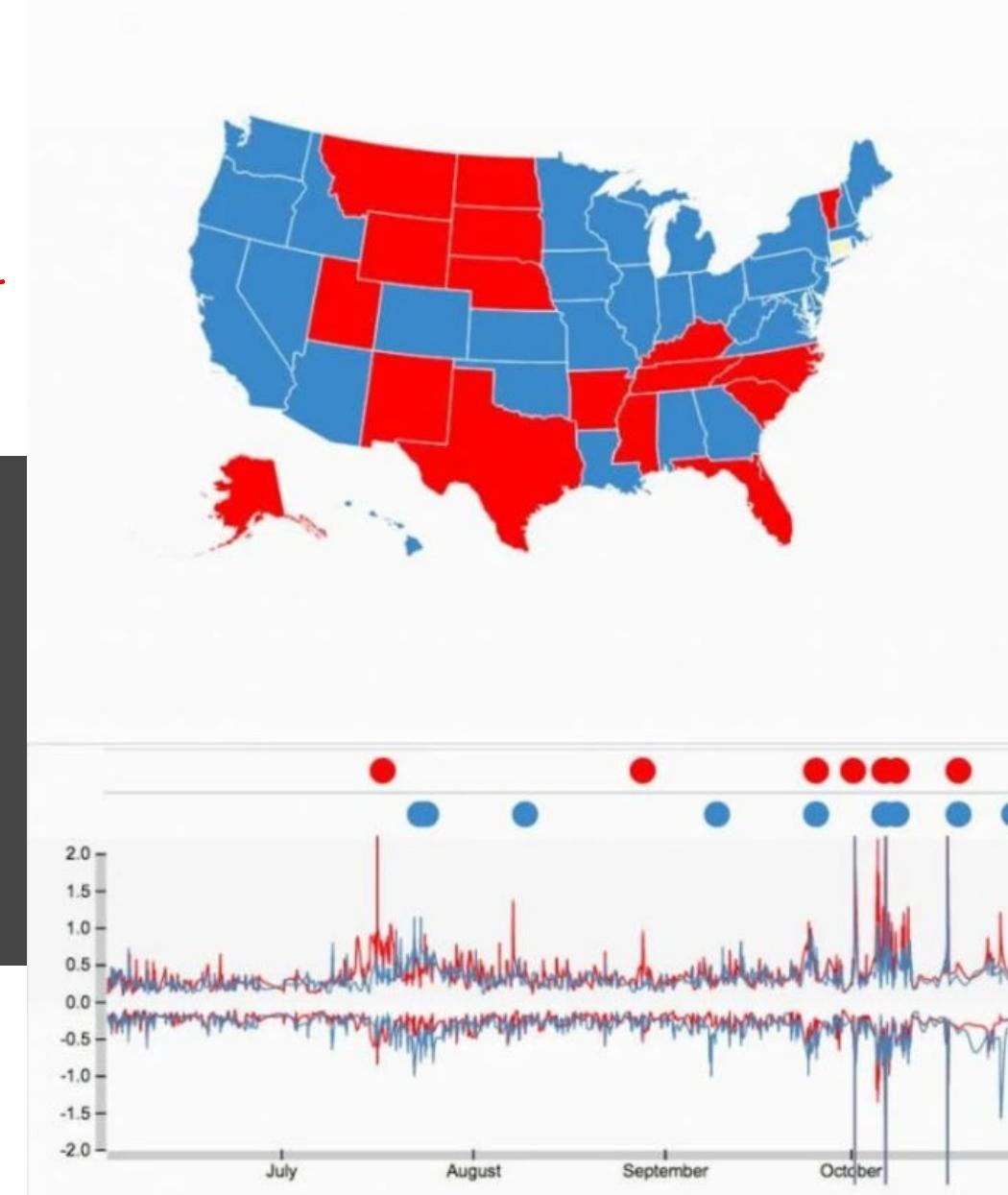
Emmy-winning US TV Shows

Police Detective TV Dramas

Critically Acclaimed Witty TV Shows

Sentiment Analysis

"NLP"



Dig deeper

Click and drag the timeline below the map or hover over the circles in "Event Timeline" to see how popular opinion changes over time.

[Learn more](#)

Legend

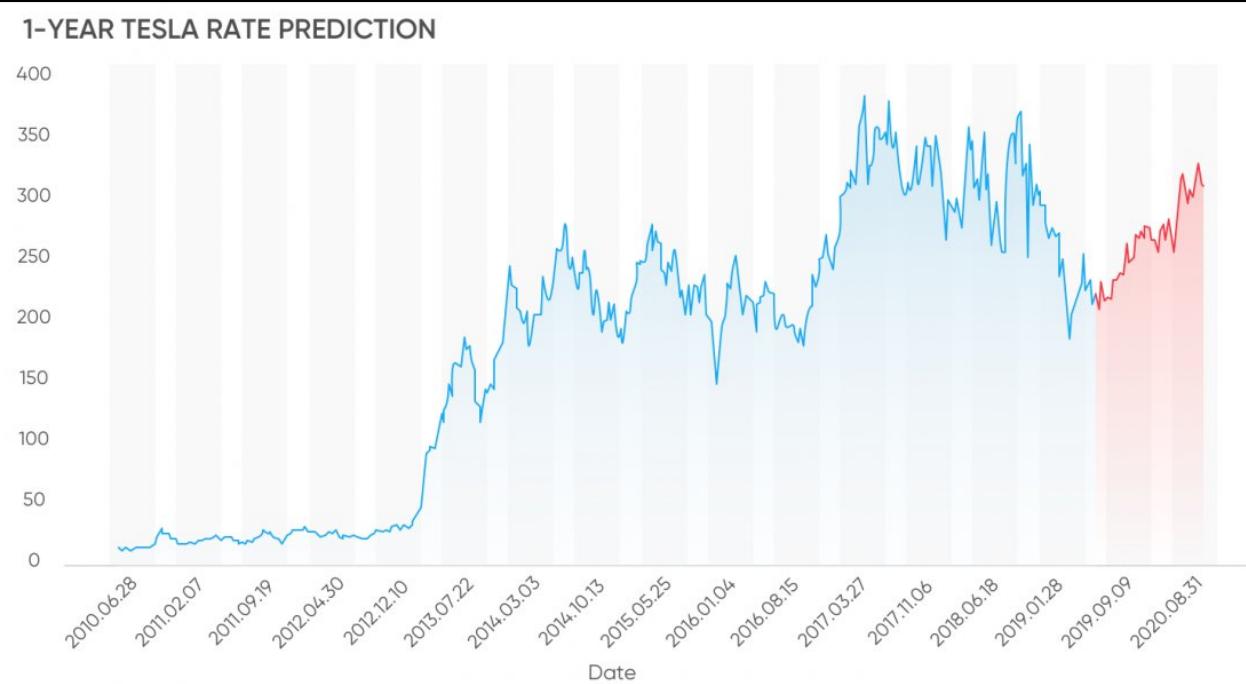
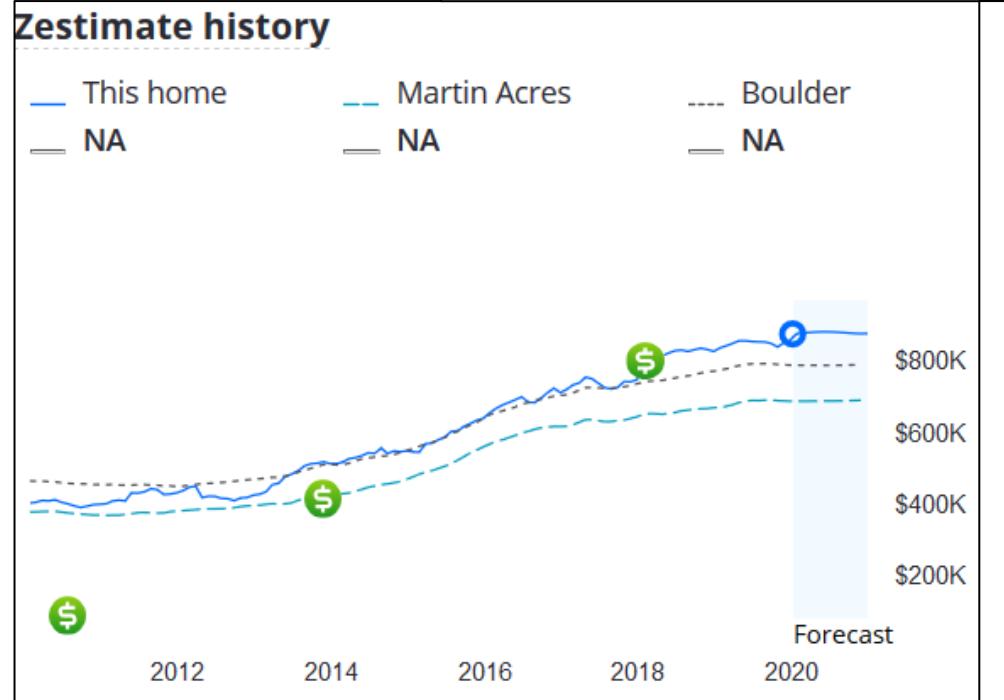
REPUBLICAN	■
DEMOCRAT	■
UNDECIDED/TIE	■
INSUFFICIENT DATA	■

Event Timeline

Positive Sentiment

Negative Sentiment

Price Forecasting



Medical Diagnosis

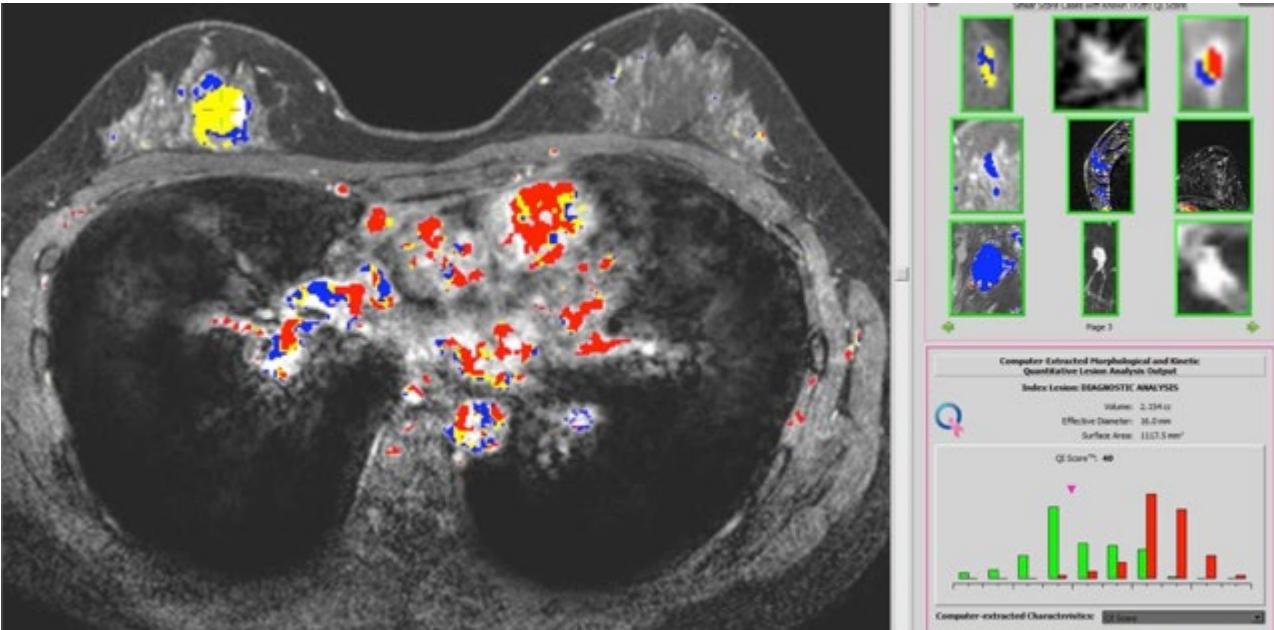
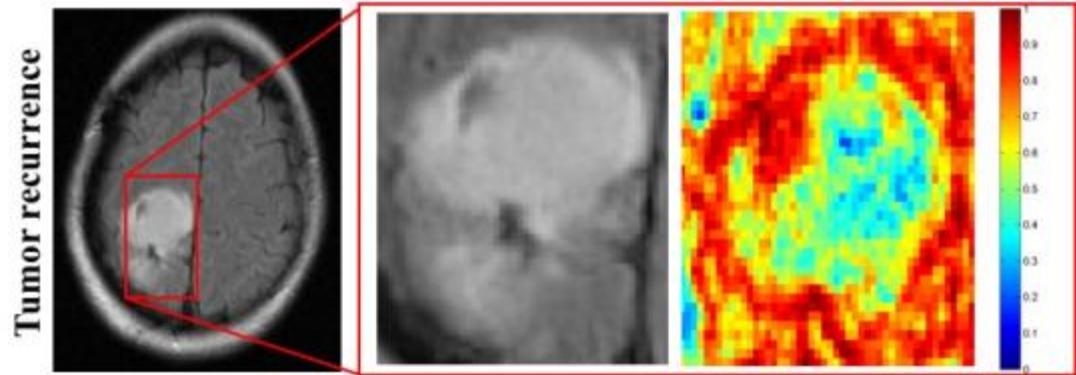
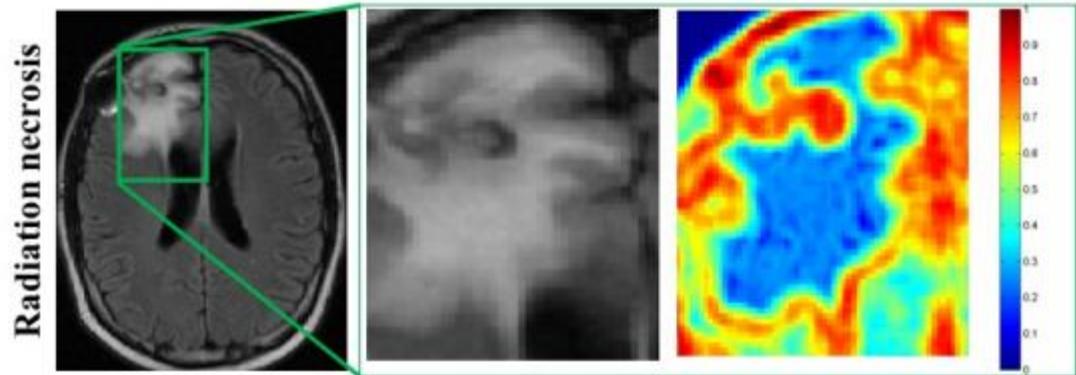
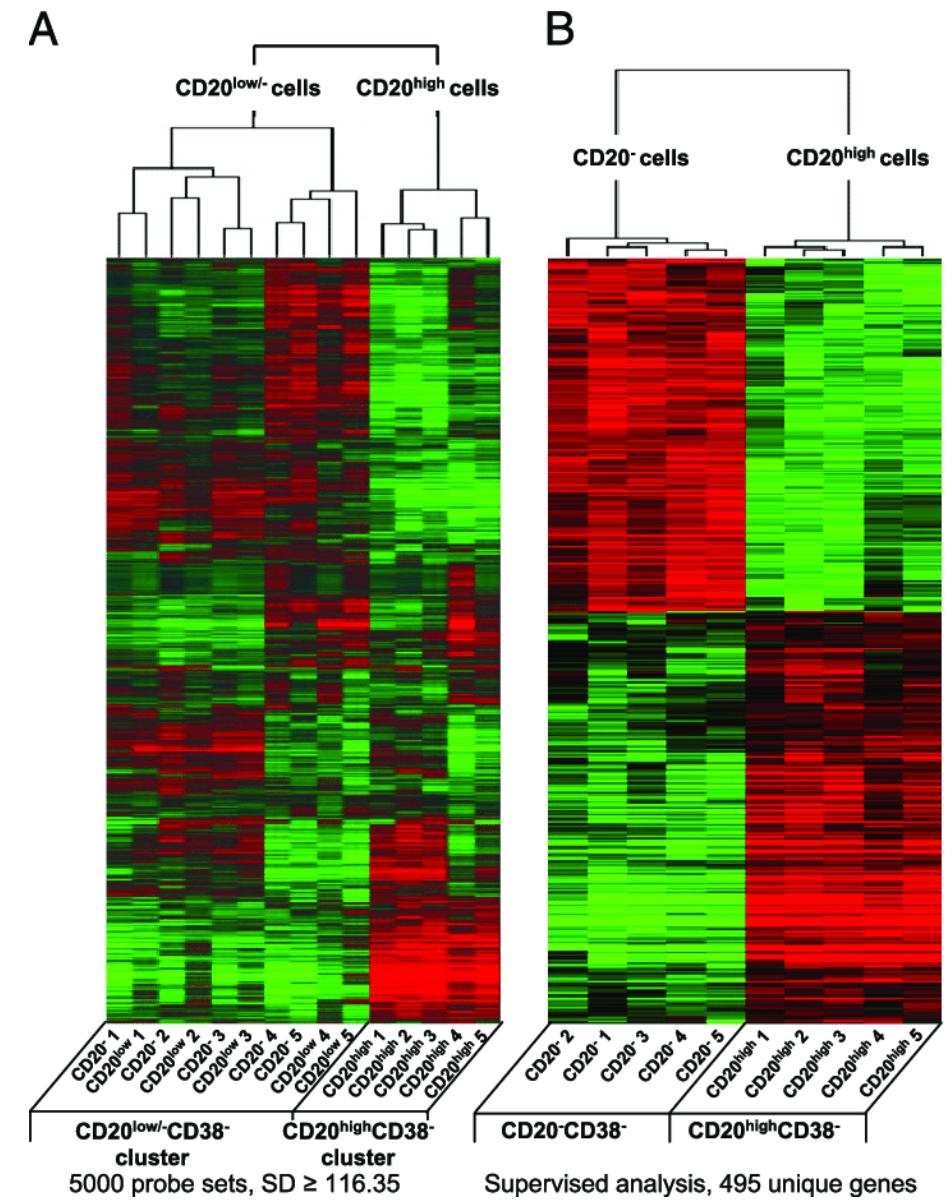
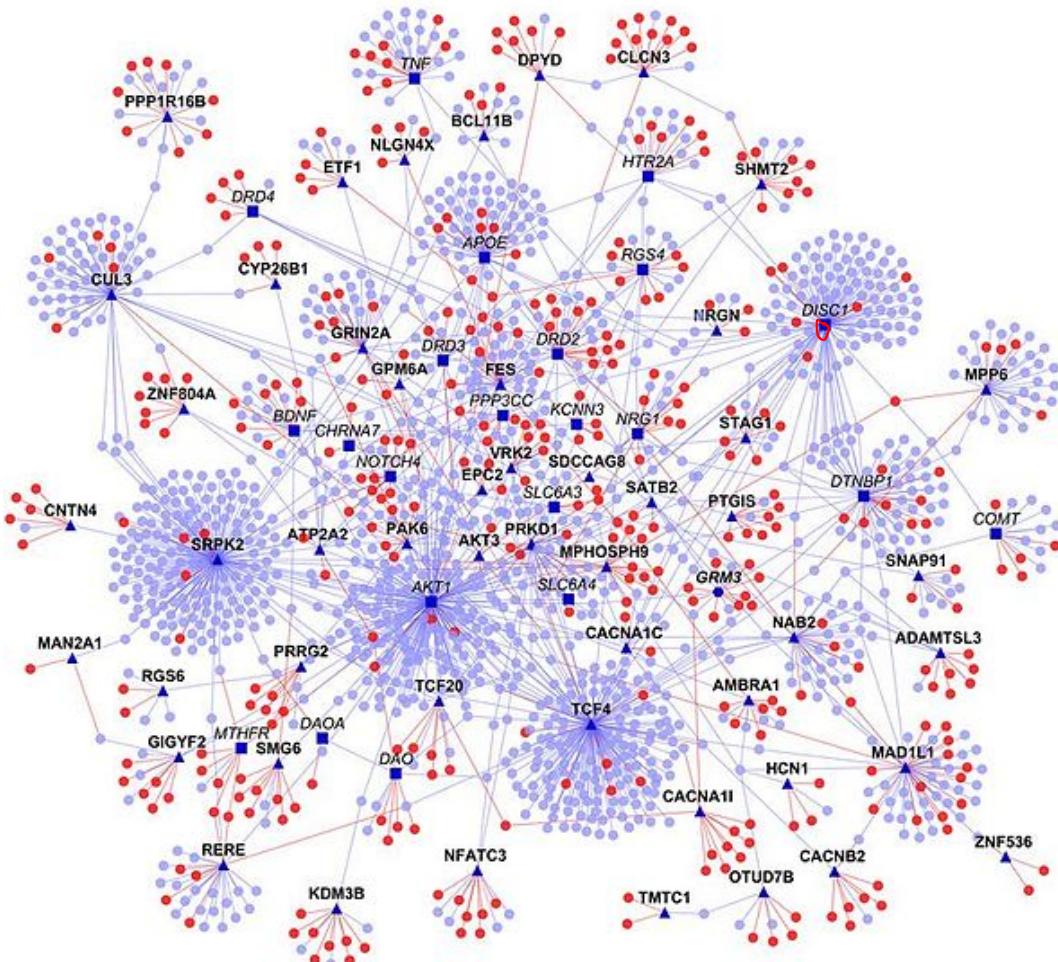


Image segmentation
(Image classification)



Bioscience Research

Unsupervised

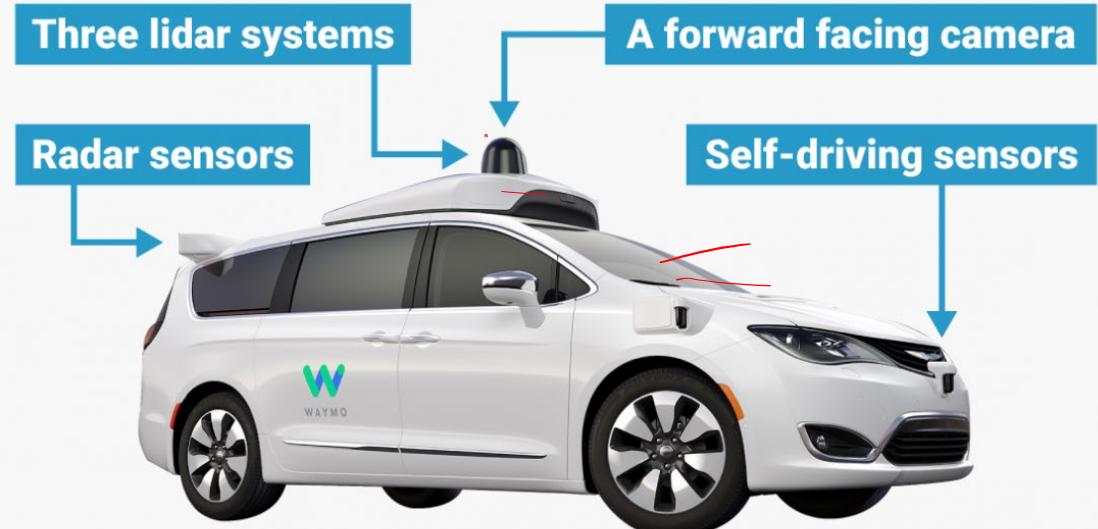


Internet of Things



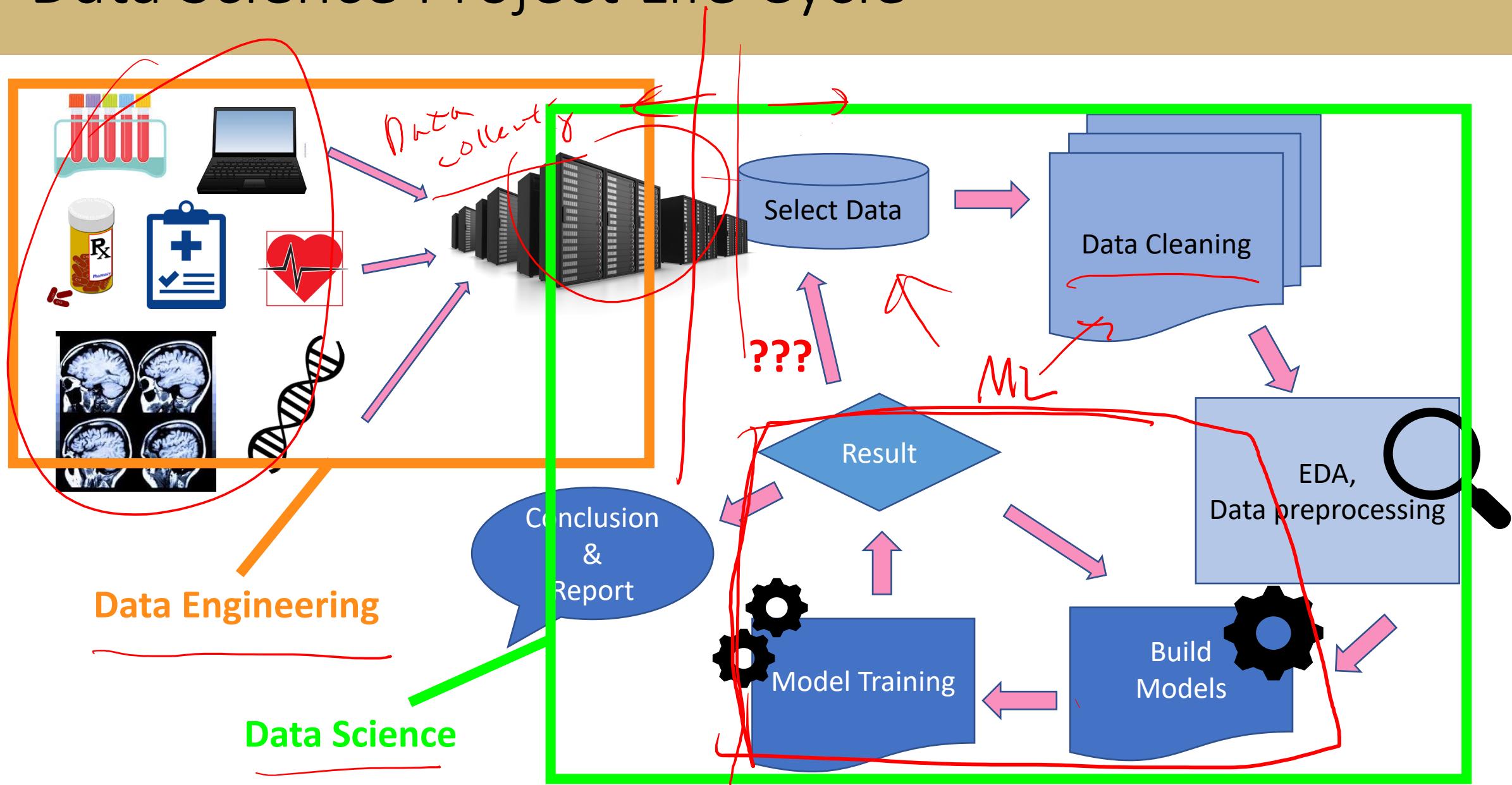
Self-driving Car

Deep Learning

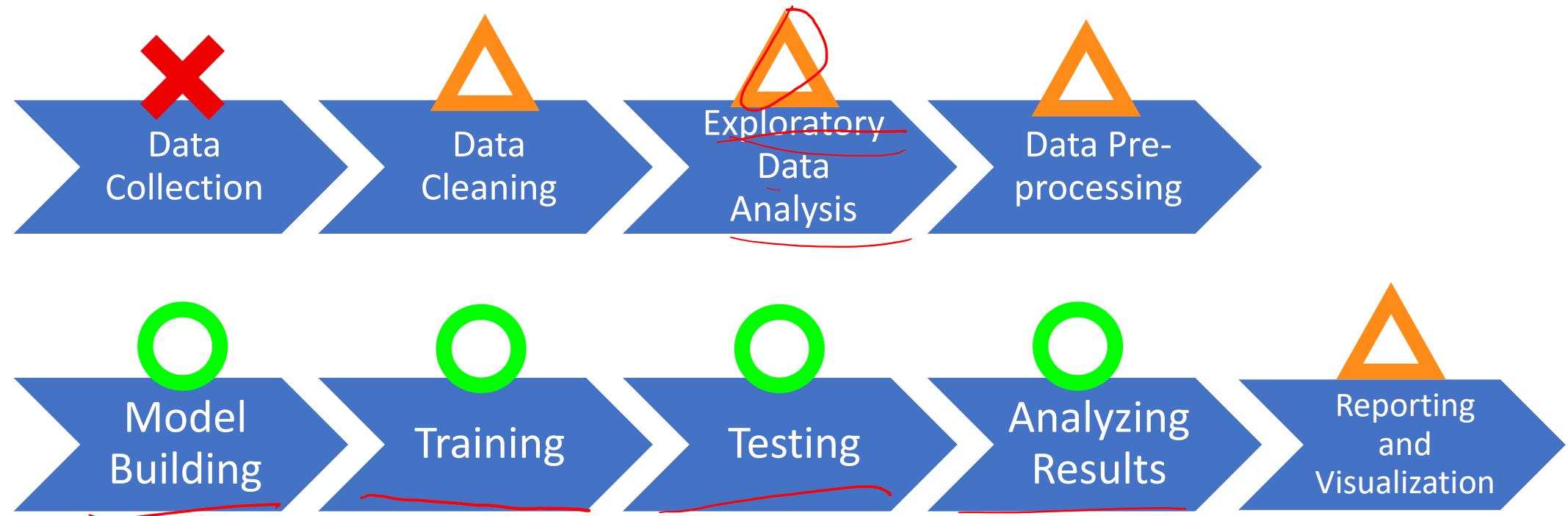


RJ

Data Science Project Life Cycle



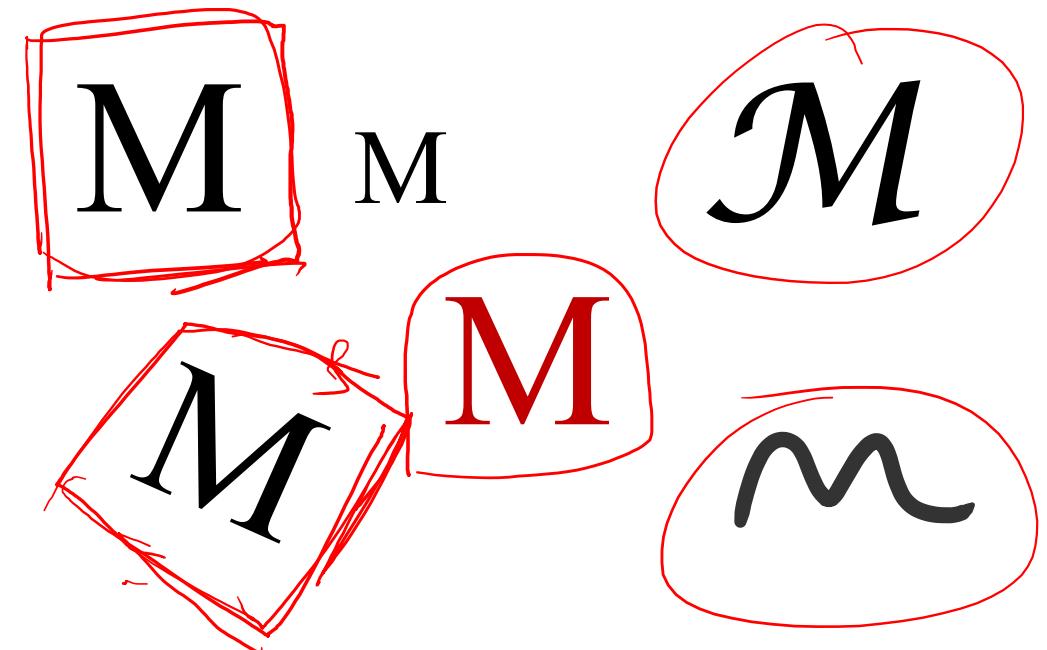
What is this course about? *ML*



What is Learning?



Learn to generalize

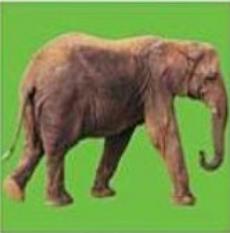


Supervised Learning

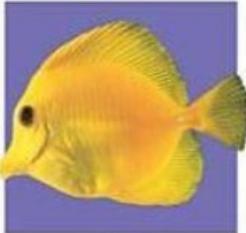
My big animal book



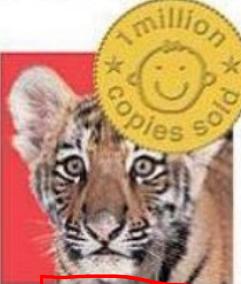
macaw



elephant



fish



tiger cub



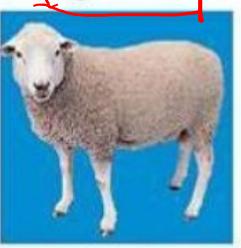
zebra



kitten



bear



sheep



duckling



dog



rabbit



cockerel



- Learn the data with labels
- Learn to predict

Unsupervised Learning

=



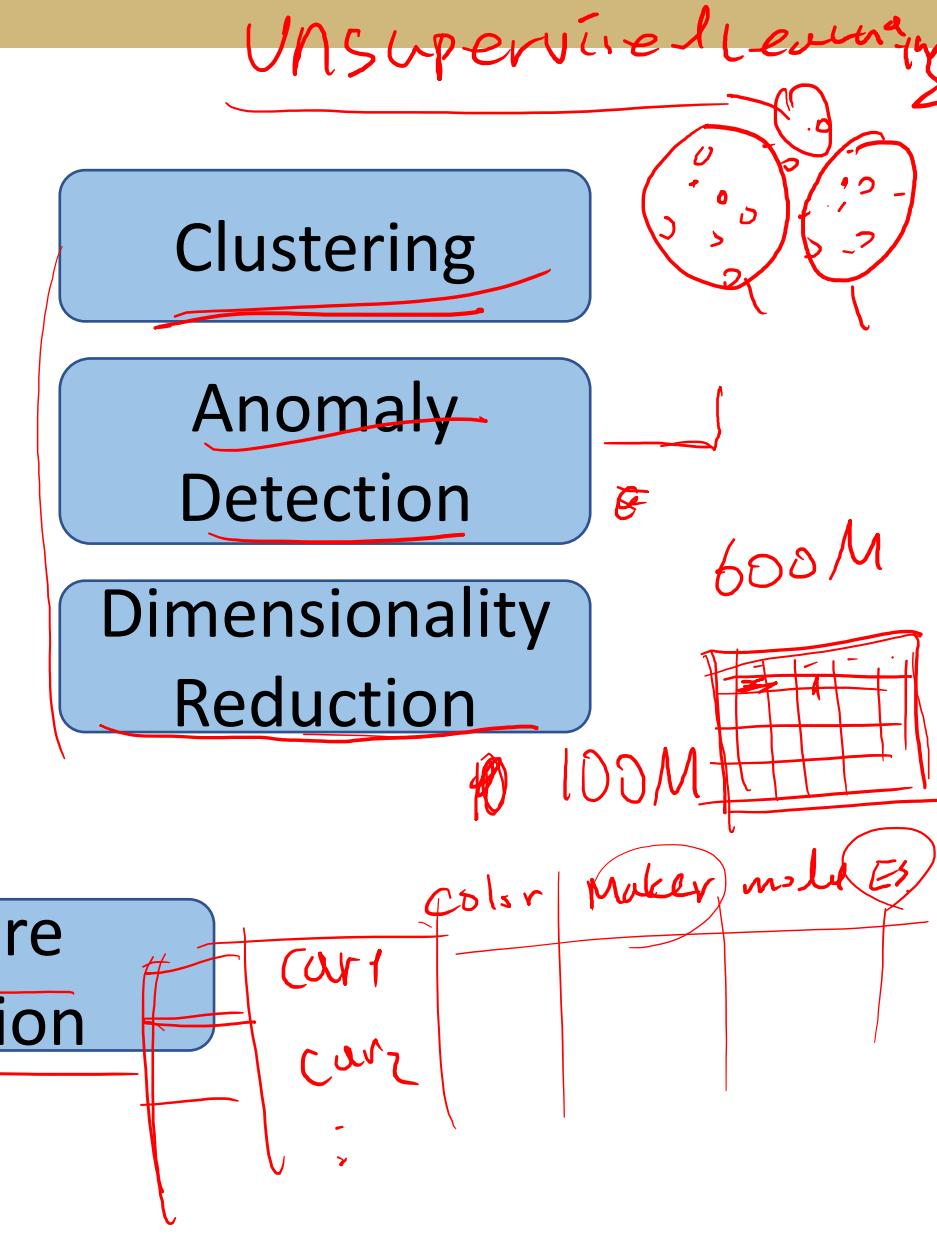
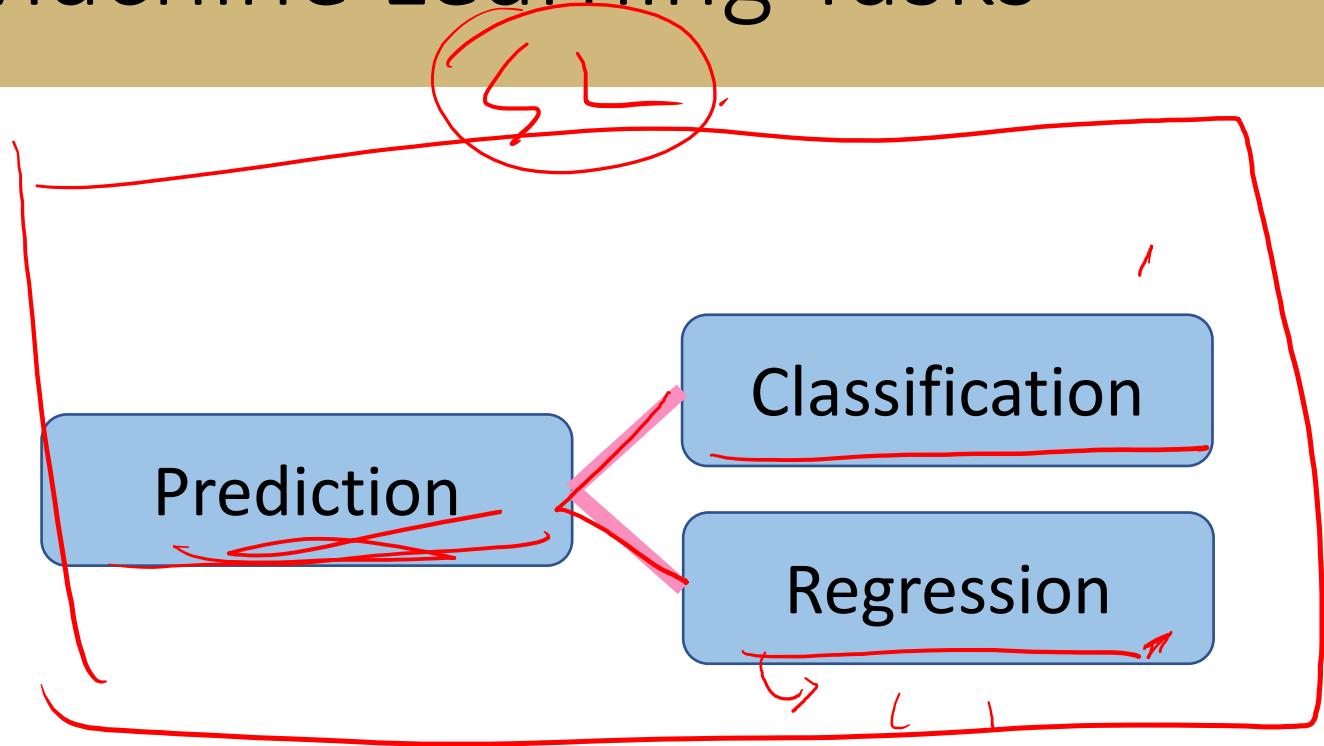
- Learn the data without labels
- Learn the underlying features/information

Reinforcement Learning

- Learn how to act from experience
- Experience = Reward/Punishment



Machine Learning Tasks



Machine Learning Models

Supervised

Parameteric

- Linear Regression (R) *regression*
- Logistic Regression (C) *classification*
- ✓ Neural Networks (R, C)
- Naïve Bayes Classifier (C)

Non-parameteric

- ✗ Support Vector Machine (R,C)
- ✗ Decision Trees (R,C)
- ✓ Nearest Neighbor (R, C)

Unsupervised

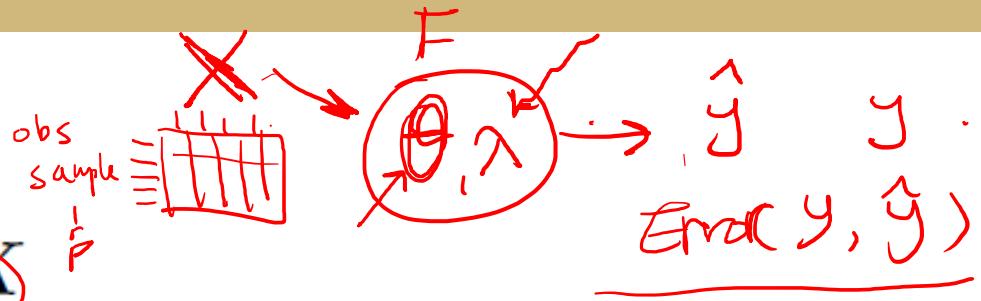
PCA (dimensionality reduction)

K-means clustering

Hierarchical clustering

Autoencoders (learning features)

Supervised Learning- variables



Data: X

Target y labels

Prediction \hat{y}

Model F

Parameters θ

Hyperparameters λ

Loss $\mathcal{L} \Rightarrow$

MSE

Cross Entropy

accuracy

Supervised Learning- model types

Data: \mathbf{X}

Target y

Prediction \hat{y}

Model \mathcal{F}

Parameters θ

Hyperparameters λ

Loss \mathcal{L}

$$\hat{y} = \mathcal{F}(\mathbf{X}) \quad \text{No } \theta, \lambda$$

$$\hat{y} = \mathcal{F}(\mathbf{X}, \underbrace{\theta_1, \theta_2, \dots, \theta_n}_n)$$

$$\hat{y} = \mathcal{F}(\mathbf{X}, \underbrace{\theta_1, \theta_2, \dots, \theta_n}_{\text{param}}, \underbrace{\lambda_1, \lambda_2, \dots, \lambda_m}_{\text{hyperparam}})$$

$$\hat{y} = \mathcal{F}(\mathbf{X}, \underbrace{\lambda_1, \lambda_2, \dots, \lambda_m}_{\text{non-parametric}})$$

DT

Supervised Learning- training

KR

LogR.

Data: \mathbf{X}

Target y

Prediction \hat{y}

Model \mathcal{F}

Parameters θ

Hyperparameters λ

Loss \mathcal{L}

$$\text{Error} = \boxed{\mathcal{L}(y, \hat{y})}$$

~~Loss~~

loss fn

Training

Optimization

MSE crossEntropy

Evaluation metric

reg: MSE, SSE.

class:

