A Story of A.I

2018, 31th January - Gameloft Indonesia @Lunch&Learn

Hello!

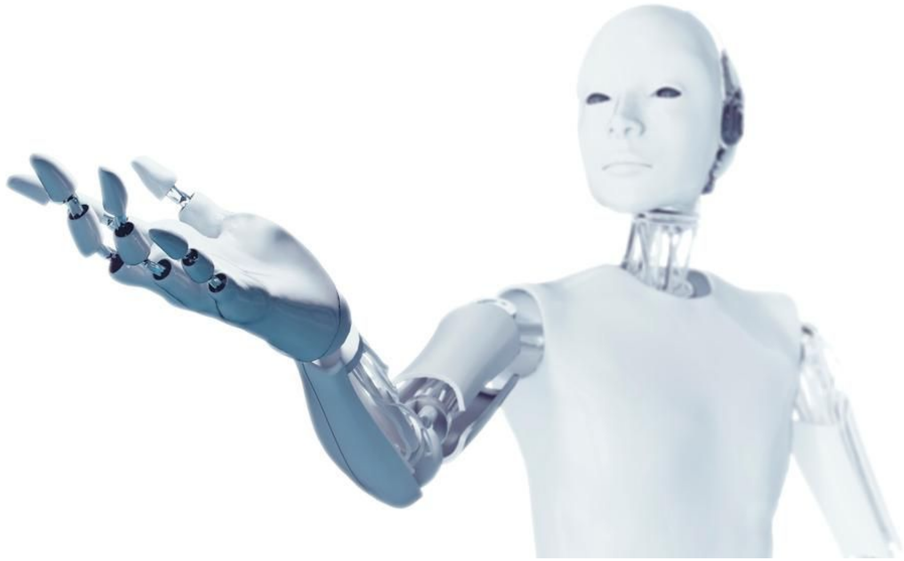
I Am Rosdyana Kusuma

- X-GL
- M.Sc Student
- Tukang Ketik Serabutan

You can contact me at
rosdyanakusuma.com

OUTLINE

- INTRODUCTION
  - A.I
  - M.L
  - D.L
- How to create simple ML / DL



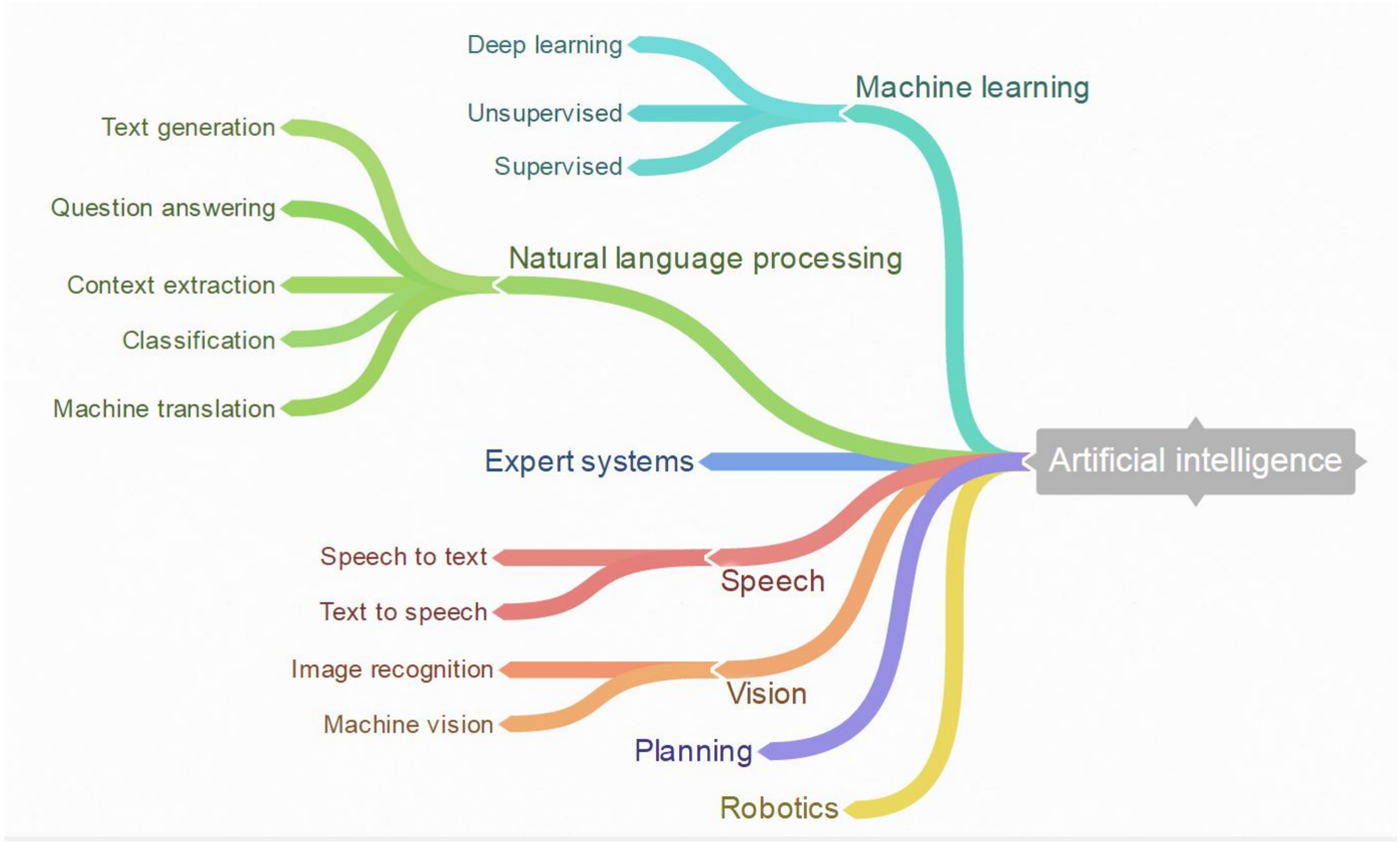
“

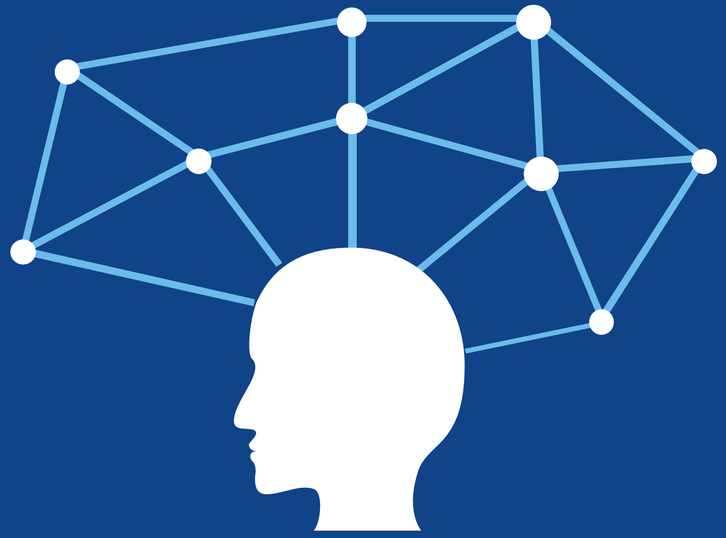
“A person who never made a mistake never tried anything new”

INTRODUCTION

What is Artificial Intelligence?

Why we should learn or approach A.I.?





Family Tree

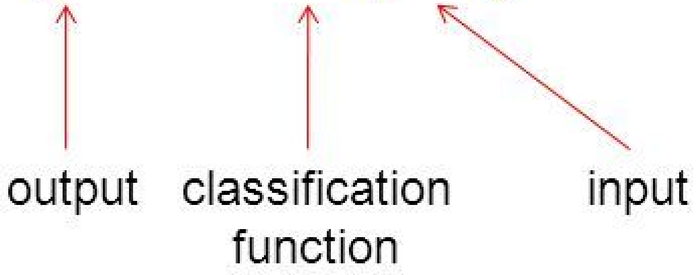
Artificial Intelligence -> Machine Learning -> Deep Learning

ML? DL?

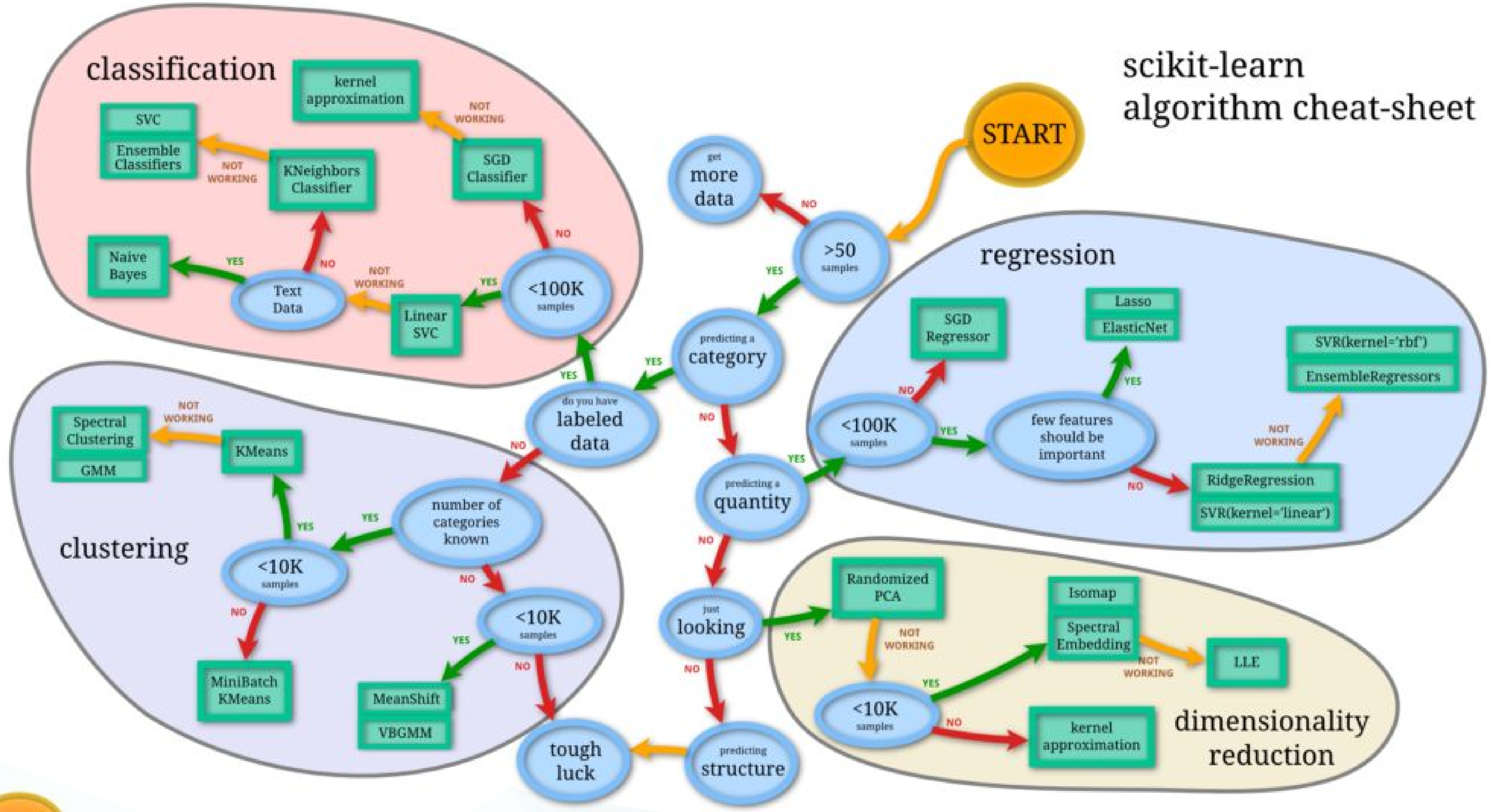
- Machine learning uses algorithms to parse data, learn from that data, and make informed decisions based on what it has learned
- Deep learning structures algorithms in layers to create an artificial “neural network” that can learn and make intelligent decisions on its own

The basic machine learning framework

$$y = f(\mathbf{x})$$



- **Learning:** given a *training set* of labeled examples $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)\}$, estimate the parameters of the prediction function f
- **Inference:** apply f to a never before seen *test example* $\mathbf{x}$ and output the predicted value $y = f(\mathbf{x})$

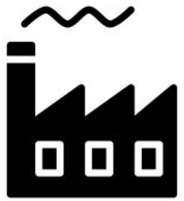

scikit-learn
algorithm cheat-sheet
START
get more data
NO
>50 samples
YES
classification
kernel approximation
NOT WORKING
SVC
Ensemble Classifiers
NOT WORKING
KNeighbors Classifier
SGD Classifier
NO
Naive Bayes
YES
NO
NOT WORKING
Text Data
NOT WORKING
Linear SVC
YES
<100K samples
NO
YES
predicting a category
YES
NO
regression
SGD Regressor
NO
Lasso
ElasticNet
YES
SVR(kernel='rbf')
EnsembleRegressors
<100K samples
YES
few features should be important
NO
NOT WORKING
RidgeRegression
SVR(kernel='linear')
do you have labeled data
NO
YES
predicting a quantity
YES
NO
clustering
Spectral Clustering
GMM
NOT WORKING
KMeans
YES
number of categories known
YES
NO
<10K samples
NO
YES
MiniBatch KMeans
<10K samples
YES
NO
MeanShift
VBGMM
just looking
NO
YES
Randomized PCA
NOT WORKING
Isomap
Spectral Embedding
NOT WORKING
LLE
<10K samples
YES
NO
kernel approximation
dimensionality reduction
tough luck
predicting structure
Back
scikit learn

REAL LIFE IMPLEMENTATION











{ML} Hello World

ML TOOLS

JAVA

- WEKA

PYTHON

- KERAS
- TENSORFLOW
- SCIKIT-LEARN
- THEANO
- MXNET

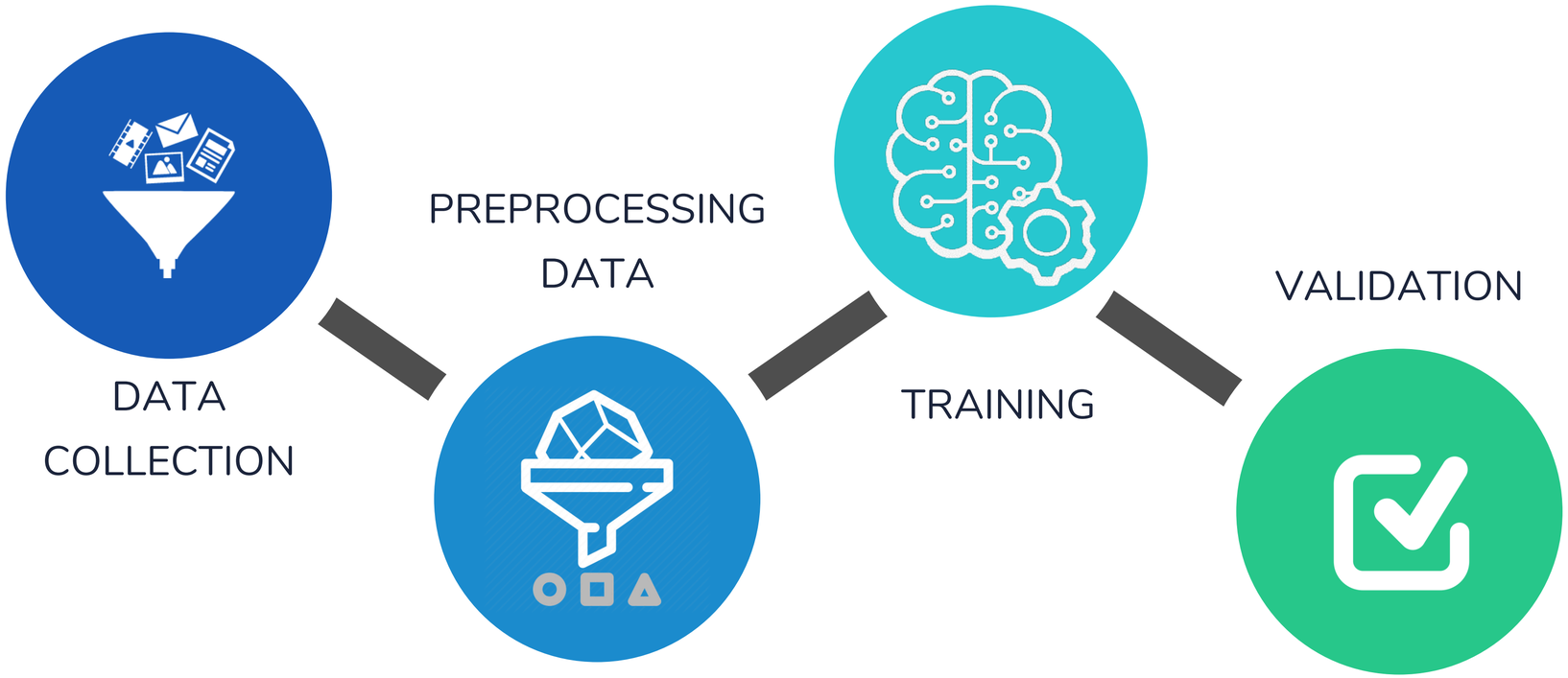
R

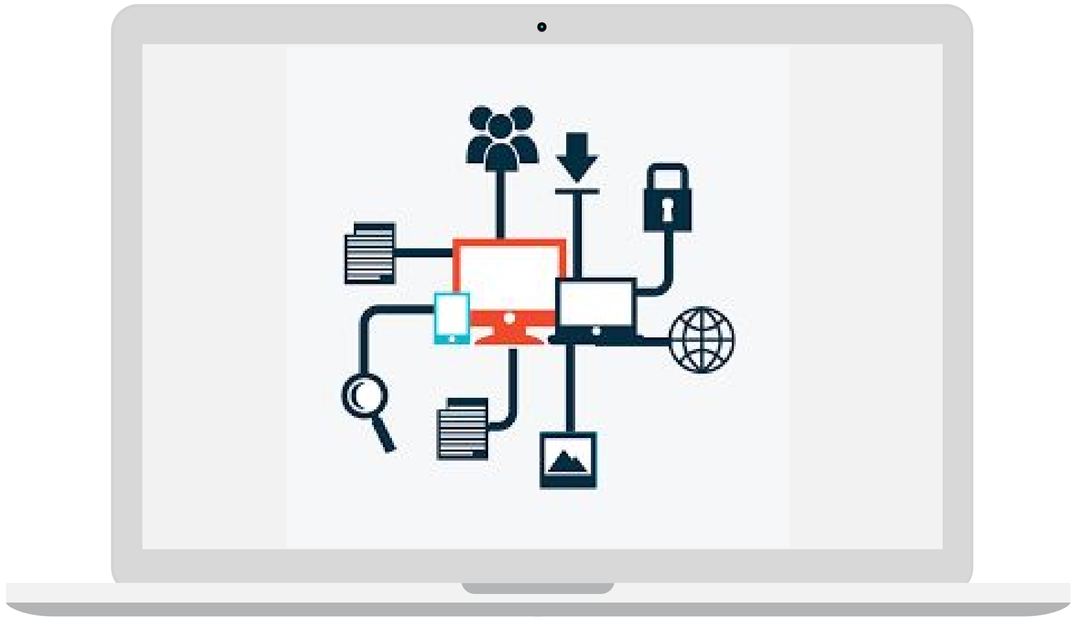
- CARET
- E1071
- TREE
- RANDOMFOREST

C++

- MLPACK

PIPELINE



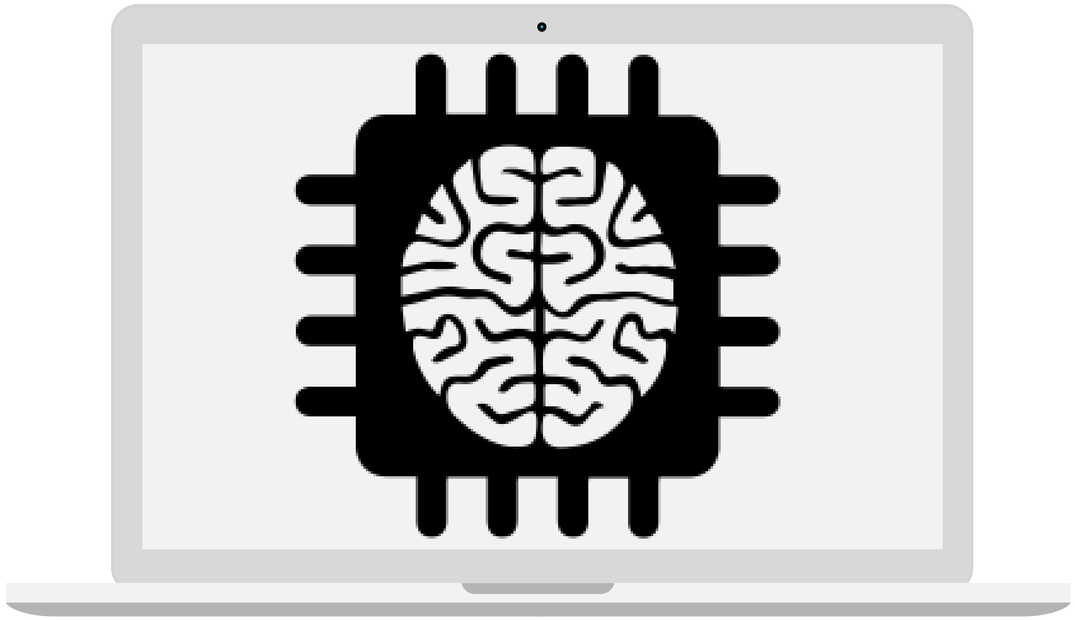


DATA COLLECTION

- https://www.kaggle.com/datasets
- http://www.uniprot.org/
-

PREPROCESSING DATA

- NORMALIZATION
- FEATURE EXTRACTION
-



TRAINING

- MACHINE LEARNING ALGO
- DEEP LEARNING ALGO
- REINFORCEMENT LEARNING ALGO

VALIDATION

- SENSITIVITY & SPECIFICITY
- ROC CURVE
- AUC CURVE
-


Thanks!
Any questions?