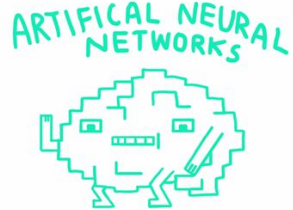
algoritma
data science education center

Torch in and

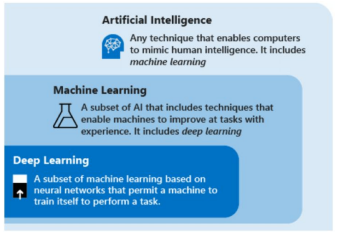
Dwi, Tomy, Wulan

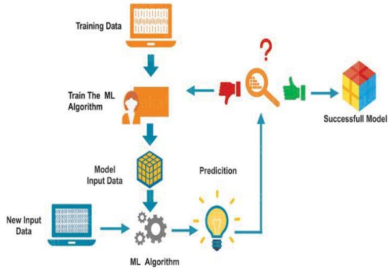


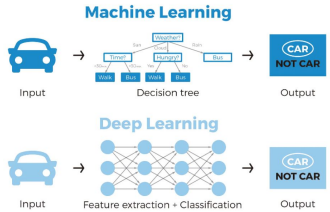
Outline

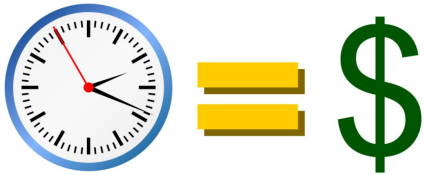
1. Background
2. Problem Statement
3. Problem Scope
4. Methodology
5. Outcome

Background

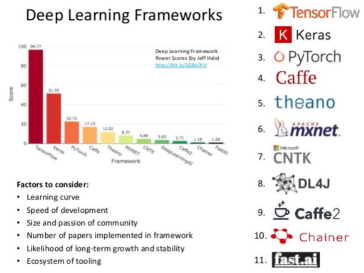

Artificial Intelligence
Any technique that enables computers to mimic human intelligence. It includes machine learning
Machine Learning
A subset of AI that includes techniques that enable machines to improve at tasks with experience. It includes deep learning
Deep Learning
A subset of machine learning based on neural networks that permit a machine to train itself to perform a task.

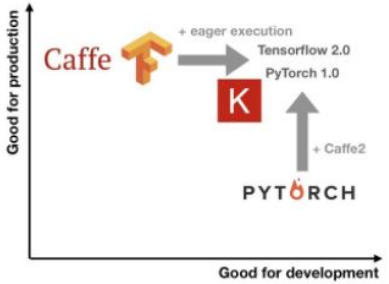

Training Data
Train The ML Algorithm
Successfull Model
Model Input Data
Predicition
New Input Data
ML Algorithm


Machine Learning
Input
Decision tree
CAR
NOT CAR
Output
Deep Learning
Input
Feature extraction + Classification
CAR
NOT CAR
Output




Chainer
mxnet
Caffe2
Microsoft CNTK
TensorFlow
Keras
GLUON
PYTORCH
theano


Deep Learning Frameworks
Factors to consider:
Learning curve
Speed of development
Size and passion of community
Number of papers implemented in framework
Likelihood of long-term growth and stability
Ecosystem of tooling
1. TensorFlow
2. Keras
3. PyTorch
4. Caffe
5. theano
6. mxnet
7. CNTK
8. DL4J
9. Caffe2
10. Chainer
11. fast.ai

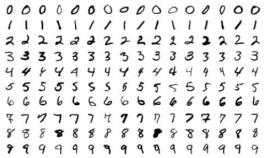

Caffe
+ eager execution
Tensorflow 2.0
PyTorch 1.0
+ Caffe2
PYTORCH
Good for production
Good for development

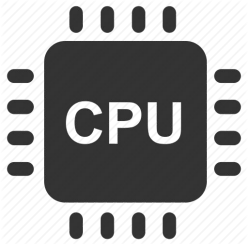
Problem Statement

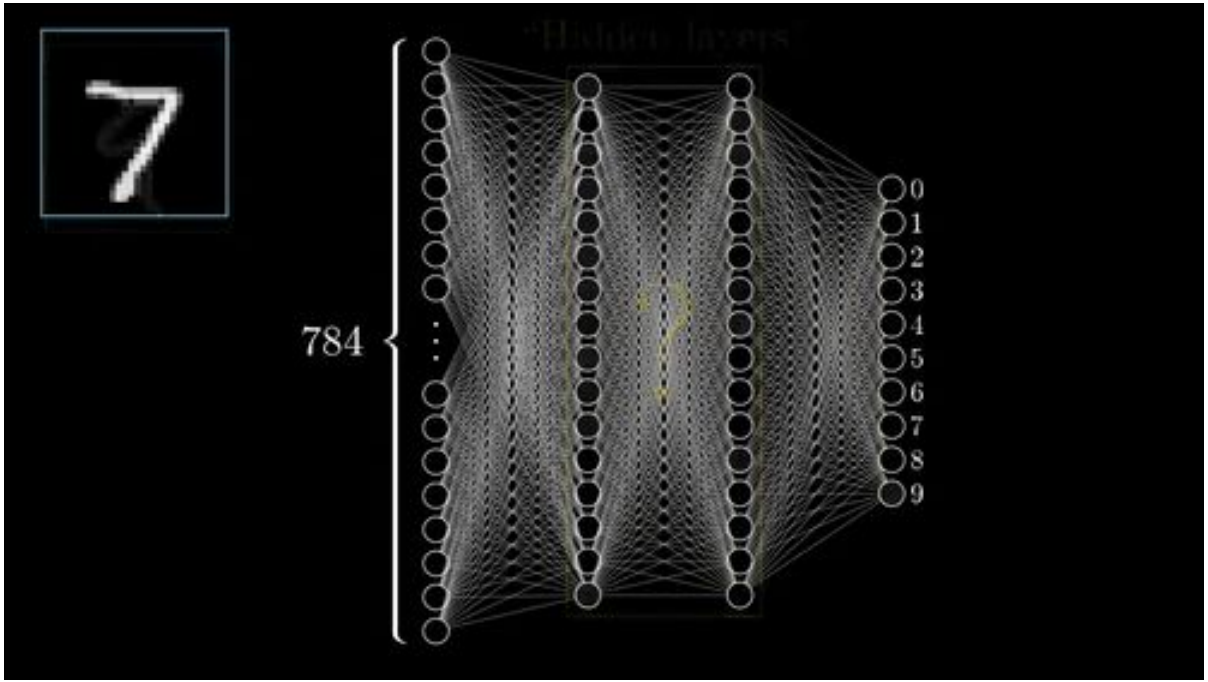
Identify digits from a dataset of handwritten images

Digit Recognizer

Problem Scope







MNIST Handwritten Digit Data
Three layers of fully-connected network

Methodology

https://pytorch.org/

https://torch.mlverse.org/

algoritma
data science education center

Load Data (Custom Dataset)

1. Read csv
2. Split target-feature
3. Tensor of pixels
4. Scaling

OOP: Inheritance concept

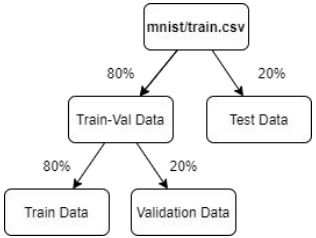
```
class MNISTDataset(Dataset):
      def __init__(self):
        ...
```

Functional

```
MNISTDataset <- dataset(
   name = ...,
   initialize = ...)
```

Splitting Data

Random splitting

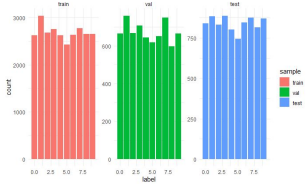


```
from torch.utils.data
import random_split
```

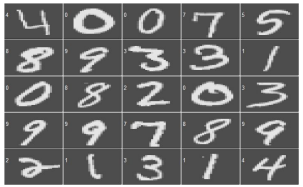
```
sample()
{base}
```

Visualize Data

Class Proportion



MNIST Data



Modelling

1. Define Model Architecture
2. Set `criterion` and `optimizer`
3. Define training and validation loop **manually**

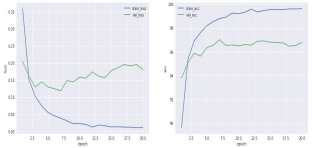
```
for e in range(n_epochs):
  ...
  for d, t in loader:
  ...
```

```
for (e in 1:n_epochs){
  coro::loop(for(b in loader){
   ...
  })
  ...
}
```

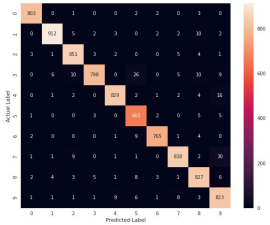
4. Save model and history performance

Evaluation

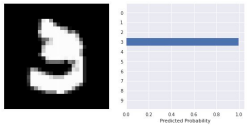
Loss and Accuracy Curves



Confusion Matrix



Unlabeled Data
mnist/test.csv



Outcome

Torch Handbook

This article contains an explanation of how to identify digits from a dataset of handwritten images using deep learning. Deep learning algorithm was developed within torch framework and implemented at Python and R.

Handbook for Python

Handbook for R