Applied Data Science and Machine Learning Course

Mohammad Sabik Irbaz
Lead Machine Learning Engineer
Pioneer Alpha Ltd.
sabik@pioneeralpha.com

Amar
iSchool

Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

Introduction

Pioneer Alpha Ltd. is an emerging R&D based Computer Software Company which provides IT-enabled products and services rendered by internet or electronic network.

MSI

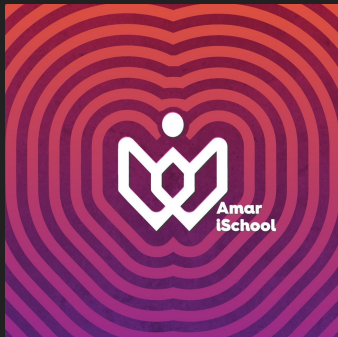# Introduction

Pioneer Alpha Ltd. is an emerging R&D based Computer Software Company which provides IT-enabled products and services rendered by internet or electronic network.

Amar iSchool is an online interactive educational platform for learners. You can increase your expertise level with Amar iSchool by watching courses and interacting with the peers and experts anytime, anywhere.

MSI

Amar iSchool

Powered by Pioneer Alpha Ltd.

"Applied Data Science and Machine Learning"
a course by Amar iSchool

Applied Data Science and Machine Learning

From the amount of data that is being produced every day, its use in business, politics, administration, or social work is increasing day by day. The arrival of 5G will further speed up the amount of this data. Then we may have to work with big data all the time. Currently, the global job growth rate in data science is around 650%. Python is like the Swiss Army Knife of technology; Python solves any task starting from the web. So if you don't enroll in this fantastic combo course in Python and Data Science, you will regret later.

BEGINNER ★★★★★ 0 (0 Ratings) 10 Students enrolled

Created by Mridul Banik Last updated Tue, 09-Feb-2021 💬 Bengali

MSI

Amar iSchool

Powered by Pioneer Alpha Ltd.

"Applied Data Science and Machine Learning" a course by Amar iSchool

Applied Data Science and Machine Learning

From the amount of data that is being produced every day, its use in business, politics, administration, or social work is increasing day by day. The arrival of 5G will further speed up the amount of this data. Then we may have to work with big data all the time. Currently, the global job growth rate in data science is around 650%. Python is like the Swiss Army Knife of technology; Python solves any task starting from the web. So if you don't enroll in this fantastic combo course in Python and Data Science, you will regret later.

BEGINNER ★★★★★ 0 (0 Ratings)   10 Students enrolled

Created by Mridul Banik   Last updated Tue, 09-Feb-2021   💬 Bengali

What will i learn?

✓ Learn Advanced Python3

✓ Data mining and wrangling

✓ How to use power BL for Data Exploration

✓ Data Modeling and Validation for supervised and unsupervised Machine Learning

✓ Deploying Predictive Models to Production

✓ Data scrapping (From web page / JSON based API)

✓ Exploratory Data analysis

✓ Data Pre-processing

✓ Data Science in Cloud

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

"Applied Data Science and Machine Learning" a course by Amar iSchool

Applied Data Science and Machine Learning

From the amount of data that is being produced every day, its use in business, politics, administration, or social work is increasing day by day. The arrival of 5G will further speed up the amount of this data. Then we may have to work with big data all the time. Currently, the global job growth rate in data science is around 650%. Python is like the Swiss Army Knife of technology; Python solves any task starting from the web. So if you don't enroll in this fantastic combo course in Python and Data Science, you will regret later.

BEGINNER ★★★★★ 0 (0 Ratings) 10 Students enrolled

Created by Mridul Banik   Last updated Tue, 09-Feb-2021   💬 Bengali

What will i learn?

✓ Learn Advanced Python3

✓ Data mining and wrangling

✓ How to use power BL for Data Exploration

✓ Data Modeling and Validation for supervised and unsupervised Machine Learning

✓ Deploying Predictive Models to Production

✓ Data scrapping (From web page / JSON based API)

✓ Exploratory Data analysis

✓ Data Pre-processing

✓ Data Science in Cloud

Curriculum for this course	3 Lessons	00:00:00 Hours
− Course Inauguration		00:00:00 Hours
+ Environment Setup	2 Lessons	00:00:00 Hours
+ Python Programming for Data Science	1 Lessons	00:00:00 Hours
+ Diving into Data Science	0 Lessons	00:00:00 Hours
+ Data Collection	0 Lessons	00:00:00 Hours
+ Diving into Data Exploration	0 Lessons	00:00:00 Hours
+ Diving into Data Preprocessing	0 Lessons	00:00:00 Hours
+ Power BI, Tableau, Introduction of KNIME	0 Lessons	00:00:00 Hours
+ Diving into Data Modeling and Validation	0 Lessons	00:00:00 Hours
+ Data Science in Cloud	0 Lessons	00:00:00 Hours
+ Deploying Predictive Models to Production	0 Lessons	00:00:00 Hours

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

"Applied Data Science and Machine Learning" a course by Amar iSchool

Applied Data Science and Machine Learning

From the amount of data that is being produced every day, its use in business, politics, administration, or social work is increasing day by day. The arrival of 5G will further speed up the amount of this data. Then we may have to work with big data all the time. Currently, the global job growth rate in data science is around 650%. Python is like the Swiss Army Knife of technology; Python solves any task starting from the web. So if you don't enroll in this fantastic combo course in Python and Data Science, you will regret later.

BEGINNER ★★★★★ 0 (0 Ratings) 10 Students enrolled

Created by Mridul Banik · Last updated Tue, 09-Feb-2021 · 💬 Bengali

What will i learn?

✓ Learn Advanced Python3

✓ Data mining and wrangling

✓ How to use power BL for Data Exploration

✓ Data Modeling and Validation for supervised and unsupervised Machine Learning

✓ Deploying Predictive Models to Production

✓ Data scrapping (From web page / JSON based API)

✓ Exploratory Data analysis

✓ Data Pre-processing

✓ Data Science in Cloud

Curriculum for this course	3 Lessons	00:00:00 Hours
− Course Inauguration		00:00:00 Hours
+ Environment Setup	2 Lessons	00:00:00 Hours
+ Python Programming for Data Science	1 Lessons	00:00:00 Hours
+ Diving into Data Science	0 Lessons	00:00:00 Hours
+ Data Collection	0 Lessons	00:00:00 Hours
+ Diving into Data Exploration	0 Lessons	00:00:00 Hours
+ Diving into Data Preprocessing	0 Lessons	00:00:00 Hours
+ Power BI, Tableau, Introduction of KNIME	0 Lessons	00:00:00 Hours
+ Diving into Data Modeling and Validation	0 Lessons	00:00:00 Hours
+ Data Science in Cloud	0 Lessons	00:00:00 Hours
+ Deploying Predictive Models to Production	0 Lessons	00:00:00 Hours

Requirements:
- Basics about Coding
- Computer Fundamentals

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

"Applied Data Science and Machine Learning" a course by Amar iSchool (cont.)

More stats:
- 50+ hours recorded classes
- 20+ hours live classes
- Documentation
- 15+ ML Algorithm based projects
- Real life project
- 15+ assignments and solve in live
- Project Deployment in Cloud
- Project Deployment using Flask
- Project API Creation
- Hyperparameter tuning on Existing GitHub Work
- Debugging
- Mock Test for Job Interview
- Personal Code Review
- Good performance lead to internship opportunities in Pioneer Alpha

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Machine Learning

According to Arthur Samuel, one of the pioneers of AI

"Machine Learning is the field of study that gives the computers the ability to learn without being explicitly programmed." [1959]

Machine Learning

**According to Arthur Samuel, one of the pioneers of AI**

"Machine Learning is the field of study that gives the computers the ability to learn without being explicitly programmed." [1959]

MSI
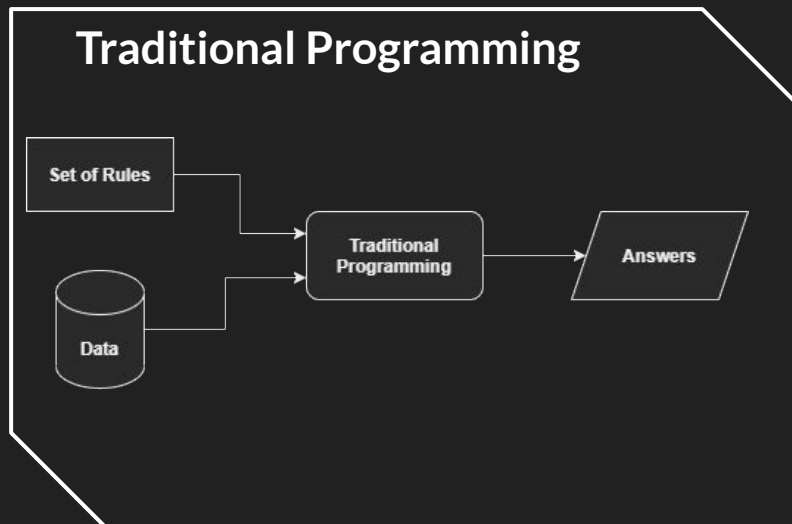
Machine Learning

According to Arthur Samuel, one of the pioneers of AI

"Machine Learning is the field of study that gives the computers the ability to learn without being explicitly programmed." [1959]

Traditional Programming

Set of Rules → Traditional Programming → Answers
Data → Traditional Programming

Machine Learning
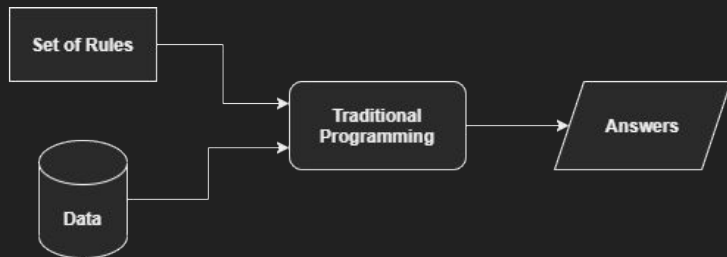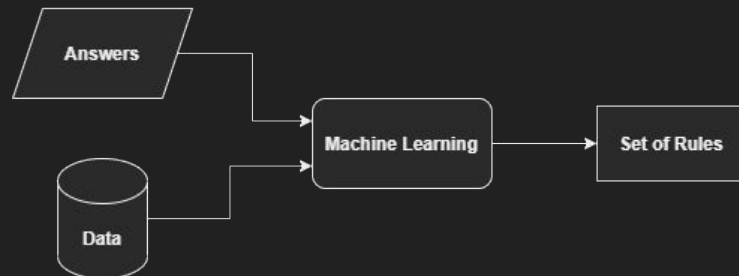
Answers → Machine Learning → Set of Rules
Data → Machine Learning

AI vs ML vs DL
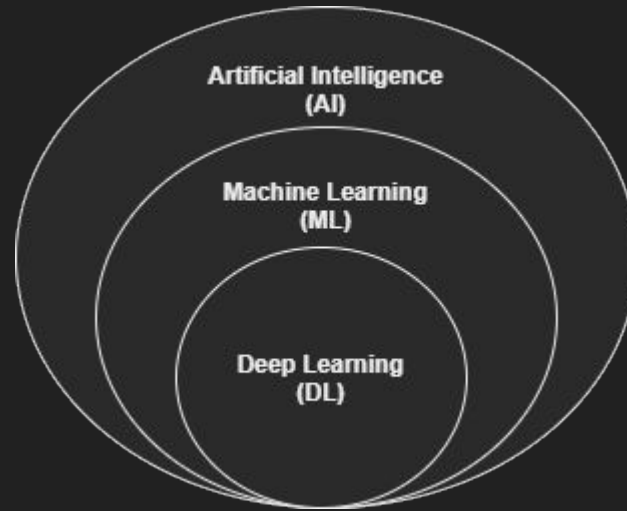
MSI

AI vs ML vs DL

Programs with the ability to learn and reason like humans

Artificial Intelligence (AI)

Machine Learning (ML)

Deep Learning (DL)
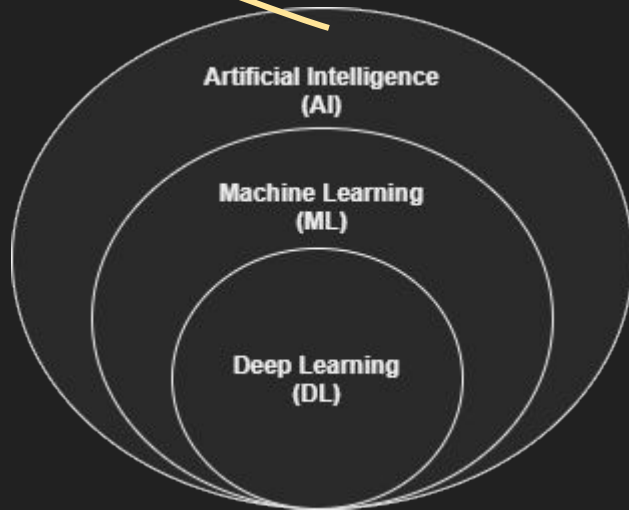
AI vs ML vs DL

Programs with the
ability to learn and
reason like humans
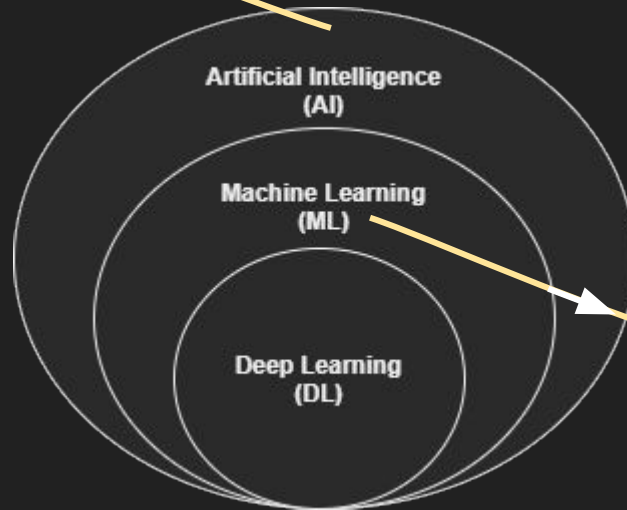
Algorithms with the
ability to learn
without being
explicitly programmed

**Artificial Intelligence
(AI)**

**Machine Learning
(ML)**

**Deep Learning
(DL)**

Subset of ML in
which Artificial
Neural Networks
adapt and learn from
vast amounts of data

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

The Origin

Arthur Samuel from IBM developed a computer program for playing checkers in the 1950s which started to beat good checkers players.

He first came up with the phrase "Machine Learning" in 1952.

Taken from Google Arts and Culture

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

The Origin (cont.)

**Perceptron
[1957]
Frank Rosenblatt combined Hebb's brain cell model with Samuel's ML efforts and created the perceptron.**

**Nearest Neighbor Algorithm
[1967]
The nearest neighbor algorithm was conceived, which was the beginning of basic pattern recognition.**
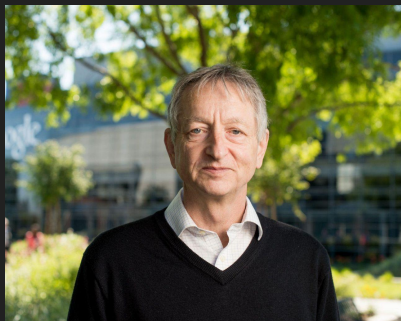
Multilayer Perceptron

There are many opinions about this. But multilayer perceptron with backpropagation was first popularized by Geoffrey Hinton.

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

Four Horsemen of Deep Learning

Geoffrey Hinton

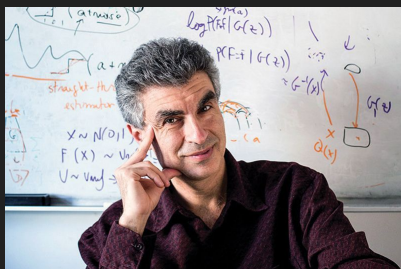invented backpropagation
won Turing Award 2018

Professor, University of Toronto
Computer scientist, Google Brain

Yann Lecun

invented convolution nets
won Turing Award 2018

VP and Chief AI Scientist,
Facebook

Yoshua Bengio

highest citation in the world in 2018
won Turing Award 2018

Co-Chair of the AI Advisory Council,
government of Canada

Andrew Ng

"democratize deep learning"
"DL with GPU revolution"

Professor, Stanford University
Founder, Google Brain

Images taken from Google Photos

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Quick Question

Can Yann LeCun sing a song?

https://twitter.com/ylecun/status/1371007395577667585
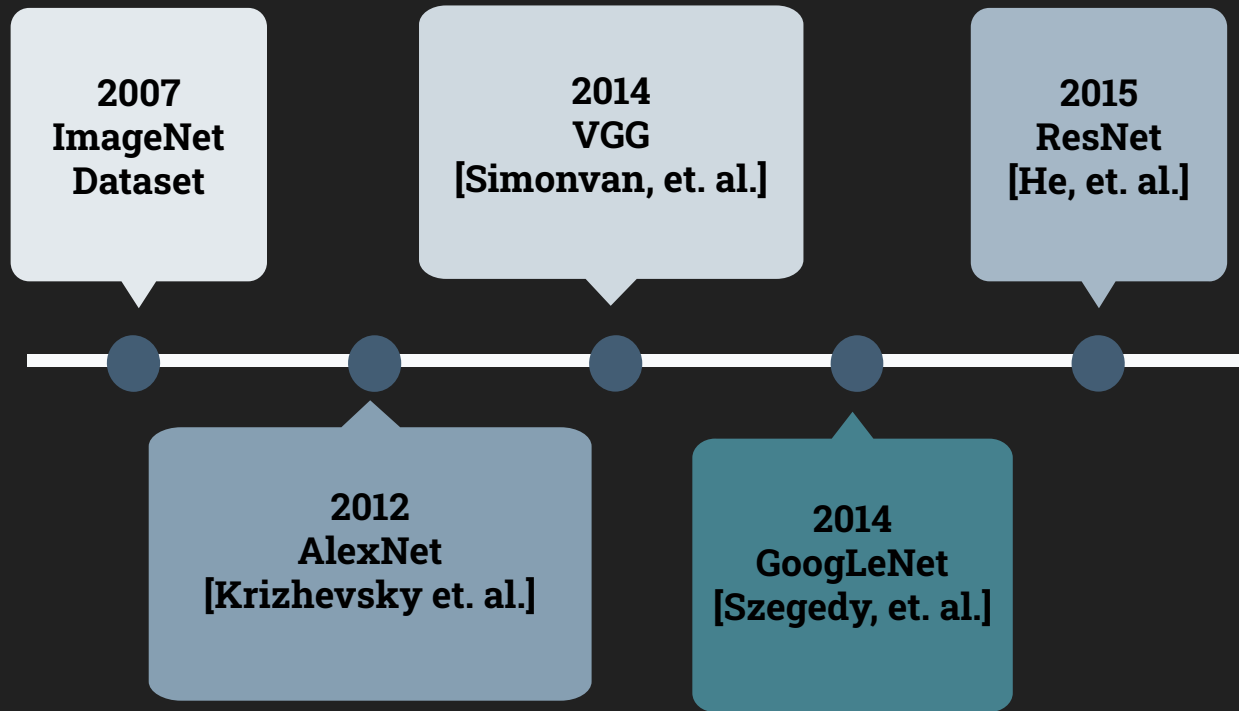
MSI

Transfer Learning

Fine-tuning pre-trained neural networks using custom layers on top for specific tasks.

Transfer Learning

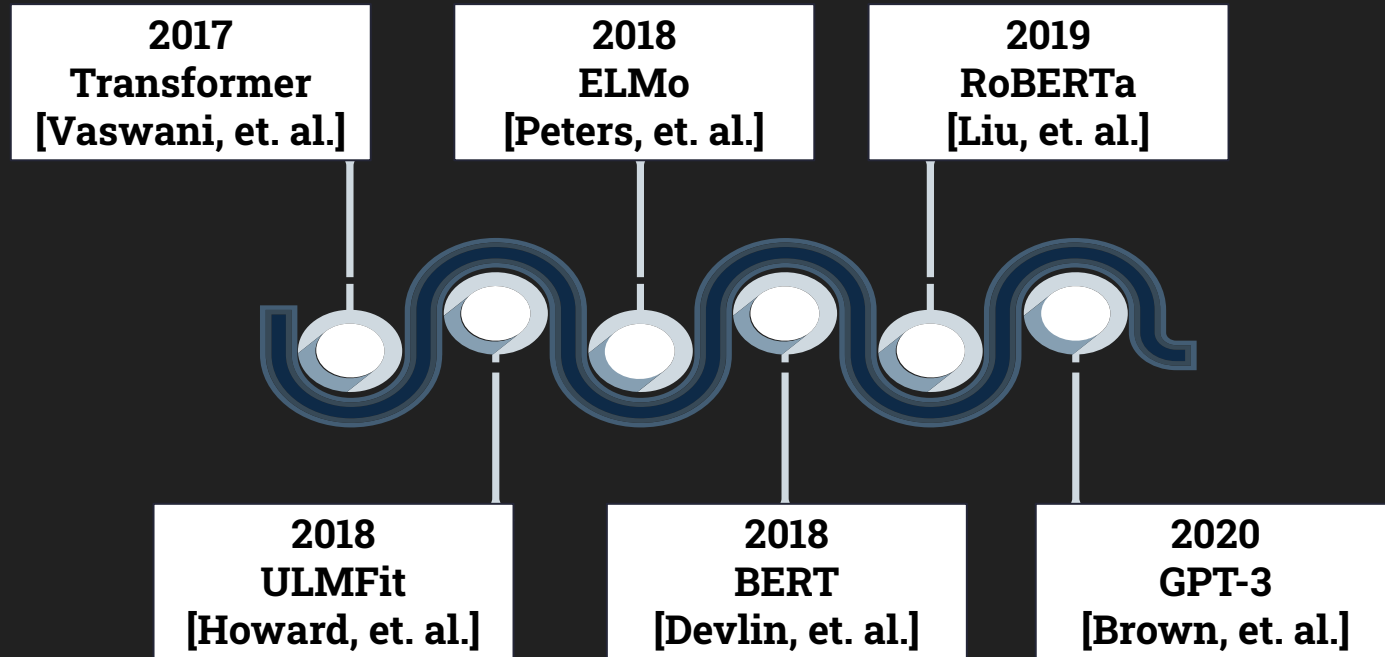Fine-tuning pre-trained neural networks using custom layers on top for specific tasks.

Images taken from Google Photos

Amar iSchool
Powered by Pioneer Alpha Ltd.

Computer Vision Revolution [2012 – 2016]

2007
ImageNet
Dataset

2014
VGG
[Simonvan, et. al.]

2015
ResNet
[He, et. al.]

2012
AlexNet
[Krizhevsky et. al.]

2014
GoogLeNet
[Szegedy, et. al.]

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

Natural Language Processing (NLP) Revolution [2017 – ongoing]

2017
Transformer
[Vaswani, et. al.]

2018
ELMo
[Peters, et. al.]

2019
RoBERTa
[Liu, et. al.]

2018
ULMFit
[Howard, et. al.]

2018
BERT
[Devlin, et. al.]

2020
GPT-3
[Brown, et. al.]

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

The Age of Transformers [ongoing]

Vision Transformer (ViT)
[Dosovitskiy et. al. 2020]

GANsformer
[Hudson et. al. 2021]

Image Segmentation
With Transformer
[Karimi et. al. 2021]

DeTr for Object Detection
[Carion et. al. 2020]

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

The Age of Transformers [ongoing]

Vision Transformer (ViT)
[Dosovitskiy et. al. 2020]

GANsformer
[Hudson et. al. 2021]

Image Segmentation
With Transformer
[Karimi et. al. 2021]

DeTr for Object Detection
[Carion et. al. 2020]

DALL-E [OpenAI 2021]

an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES

MSI

**Let's have a look at what Tom
Cruise wants to say us.**

https://twitter.com/mckaywrigley/status/1364972080513515520

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Let's have a look at what Tom
Cruise wants to say us.

https://twitter.com/mckaywrigley/status/1364972080513515520

Is AI getting dangerous?

https://ai.facebook.com/blog/deepfake-detection-challenge-resul
ts-an-open-initiative-to-advance-ai/

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Images taken from Google Photos

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

What do the experts say?

Abhisek Thakur
4x Kaggle Grandmaster

"We should always try baseline models first."

https://twitter.com/abhi1thakur/status/1369719373464748040

Jeremy Howard
Founder of fast.ai

"Baseline models work better than
your fancy models most of the time."

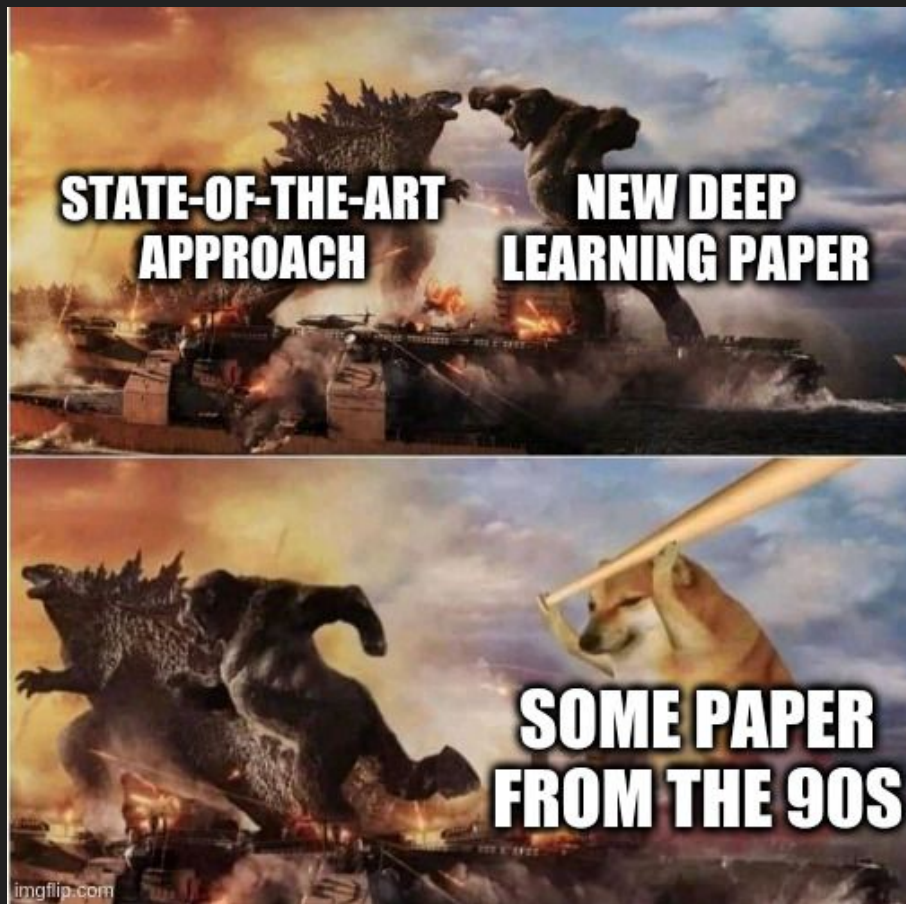https://twitter.com/omarsar0/status/1370542107232964608

MSI

Images taken from Google Photos

Amar
iSchool
Powered by Pioneer Alpha Ltd.

This is still happening

STATE-OF-THE-ART APPROACH

NEW DEEP LEARNING PAPER

SOME PAPER FROM THE 90S

imgflip.com

MSI

Images taken from Google Photos

Amar iSchool
Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

ML Production

Chip Huyen (Teaches MLSys at Stanford University)

ML in production: expectation

1. Collect data
2. Train model
3. Deploy model
4.

MSI

ML Production

Chip Huyen (Teaches MLSys at Stanford University)

ML in production: expectation

1. Collect data
2. Train model
3. Deploy model
4.

ML in production: reality

1. Choose a metric to optimize
2. Collect data
3. Train model
4. Realize many labels are wrong -> relabel data
5. Train model
6. Model performs poorly on one class -> collect more data for that class
7. Train model
8. Model performs poorly on most recent data -> collect more recent data
9. Train model
10. Deploy model
11. Dream about $$$
12. Wake up at 2am to complaints that model biases against one group
    -> revert to older version
13. Get more data, train more, do more testing
14. Deploy model
15. Pray
16. Model performs well but revenue not increases -> choose a different metric
17. Cry
18. Start over

MSI

Data Science

Images taken from Google Photos

TABLE OF CONTENTS

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Types of ML Algorithms

➔ Supervised Learning

➔ Unsupervised Learning

➔ Semi-Supervised Learning

➔ Self-Supervised Learning (SSL)

➔ Reinforcement Learning (RL)

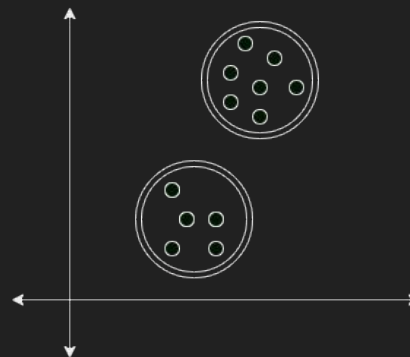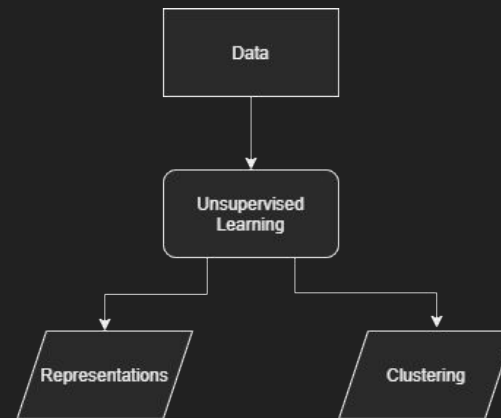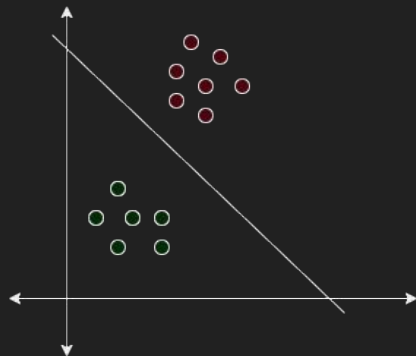➔ Active Learning
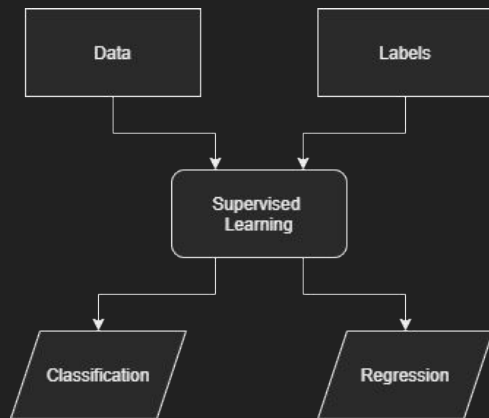
MSI

Supervised vs Unsupervised Learning

MSI

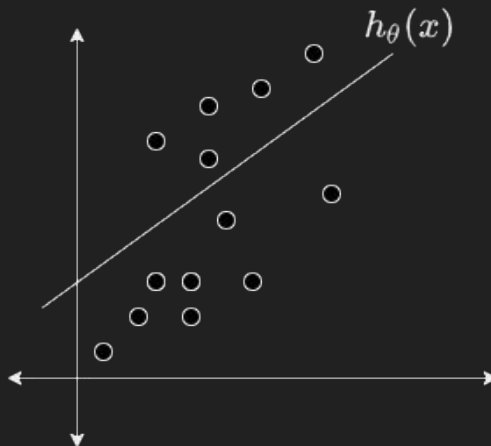# TABLE OF CONTENTS

MSI

Amar iSchool

Hypothesis
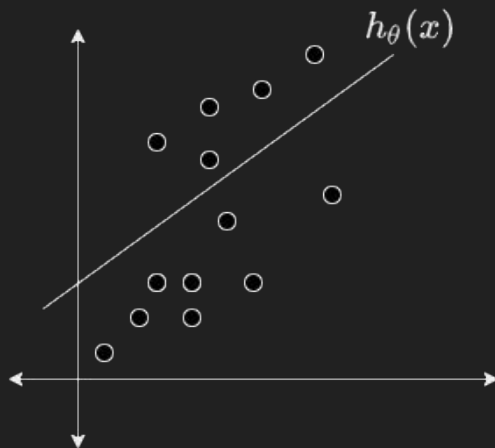
Representation of a line : y = mx + c

MSI

Hypothesis

Representation of a line : y = mx + c

Let's redefine it

$$h_\theta(x) = \theta_0 + \theta_1 * x$$

MSI

Mean Squared Error (MSE)

$h_\theta(x)$

=> Imagine you have 'm' data points

=> You drew a line

=> MSE calculates how far is that line from data points in average.

=> Mean is for the average

=> Square is for the Euclidean Distance

MSI

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Cost Function

Cost Function in case of Linear Regression is nothing but MSE

$$J(\theta_0, \theta_1) = \frac{1}{2m} * \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

m = #samples

$$h_\theta(x) = \theta_0 + \theta_1 * x$$

MSI

Cost Function

Cost Function in case of Linear Regression is nothing but MSE

$$J(\theta_0, \theta_1) = \frac{1}{2m} * \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

m = #samples

$$h_\theta(x) = \theta_0 + \theta_1 * x$$

Our Objective is to minimize the cost.

How do we do that?

MSI

Gradient Descent

Images taken from Google Photos

MSI

Gradient Descent (cont.)

Visualization by
Lili Jiang, Head of Data Science, Quora

Amar
iSchool
Powered by Pioneer Alpha Ltd.

Gradient Descent (cont.)

Repeat until Convergence

$$\theta_j = \theta_j - \alpha * \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

MSI

Gradient Descent (cont.)

Repeat until Convergence

$$\theta_j = \theta_j - \boxed{\alpha} * \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

Learning Rate

MSI

Gradient Descent (cont.)

Repeat until Convergence

$$\theta_j = \theta_j - \alpha * \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

Learning Rate

Partial Derivative of the Cost Function

MSI

Code Walkthrough

Amar iSchool
Powered by Pioneer Alpha Ltd.

TABLE OF CONTENTS

MSI

Amar iSchool
Powered by Pioneer Alpha Ltd.

References

1. Kiaming He, et. al. "Deep residual learning for image recognition". in CVPR 2016.
2. Jacob Devlin, Ming-Wei Chang. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". in NAACL-HLT 2018.
3. Ashish Vaswani, Noam Shazeer. "Attention is all you need". in NeurIPS 2017
4. Chip Huyen. "Machine Learning Systems Design"
5. Andrew Ng. "Machine Learning"
6. Fei-Fei Li. "Convolutional Neural Networks for Visual Recognition"
7. Chris Manning. "Natural Language Processing with Deep Learning"