Deep Learning for Humanists @DHSI2023 – Day 1

Hoyeol Kim

These slides will be shared on GitHub

Hoyeol Kim

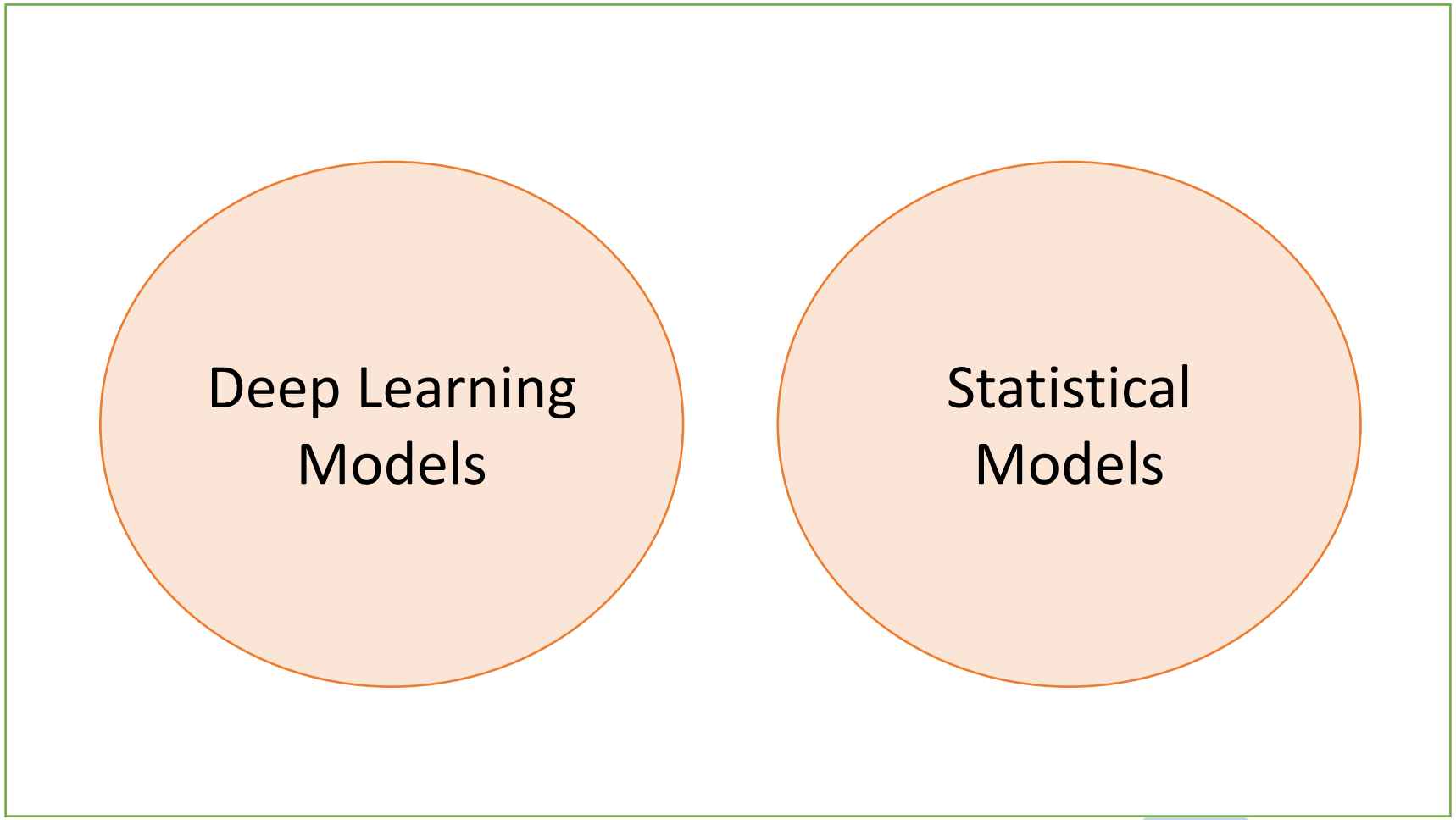
- Machine Learning Data Linguist on the AI Data team of Amazon Web Services
- PhD with a focus on computational approaches in the humanities (CV & NLP) at Texas A&M University.
- Twitter / Mastodon : @elibooklover / @hoyeolkim@hcommons.social

What are we going to learn?

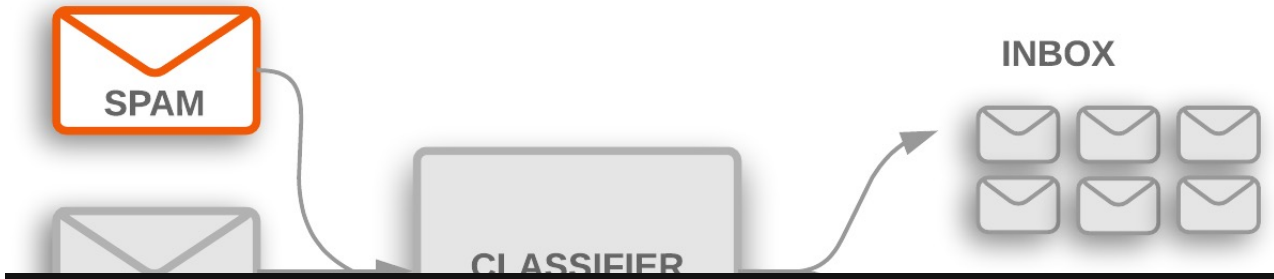
Machine Learning

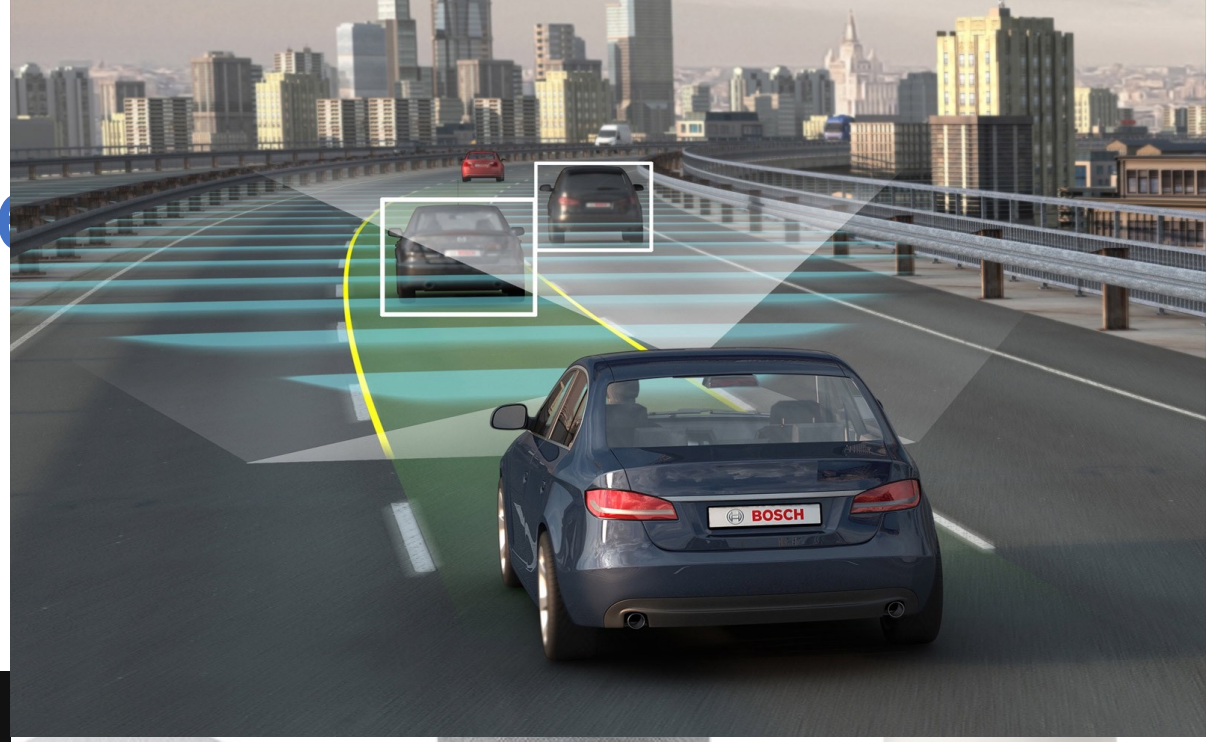
- Machine learning is a set of techniques to make computers better perform things humans can do such as calculation, memory, and transmission.
- Machine learning involves ‘learning’ as humans do.

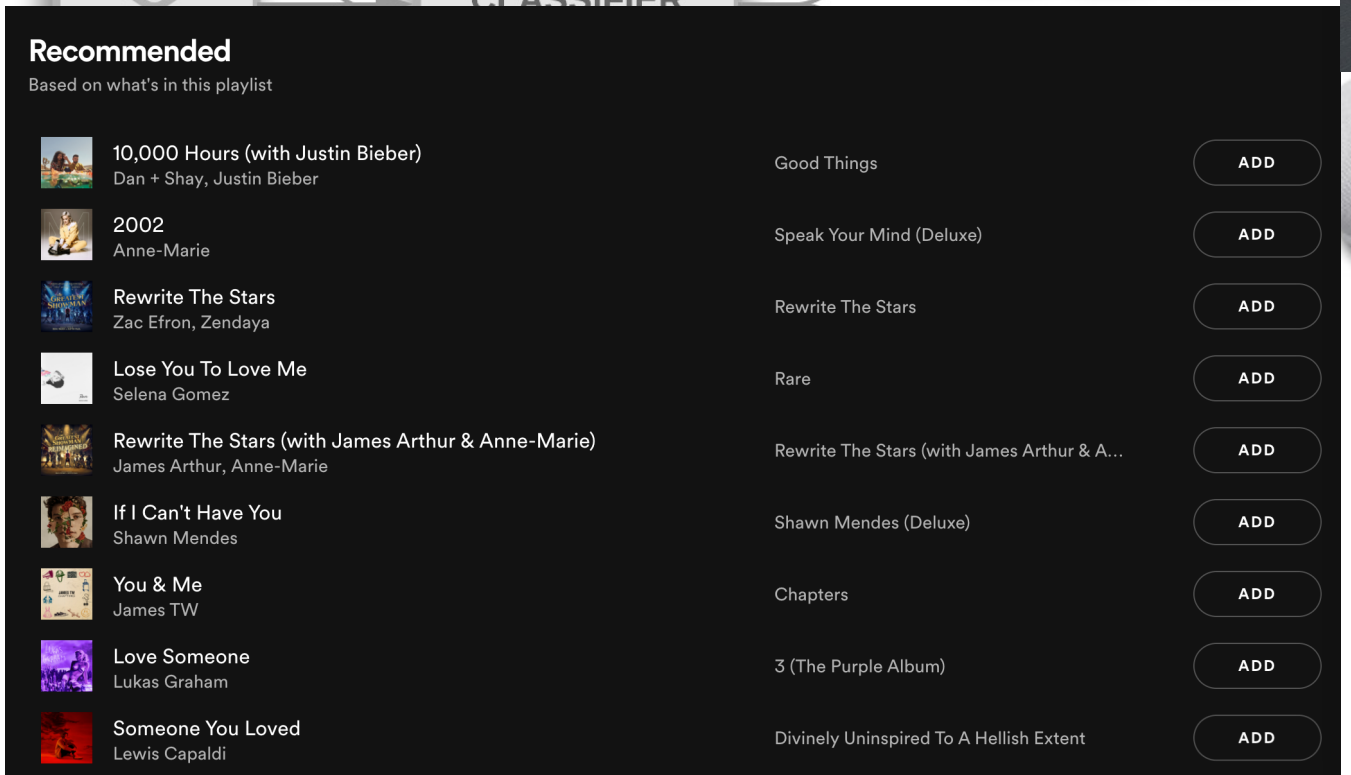
Machine Learning



Machine Le

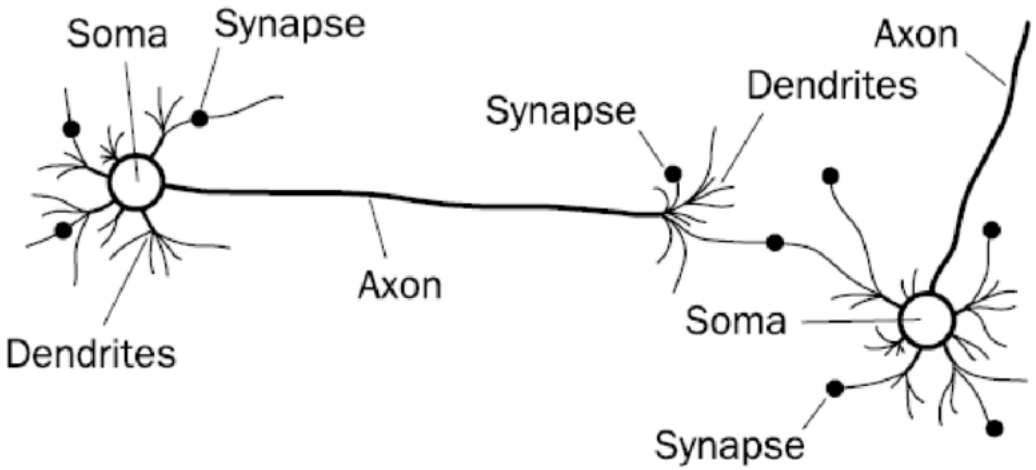

SPAM
CLASSIFIER
INBOX


BOSCH

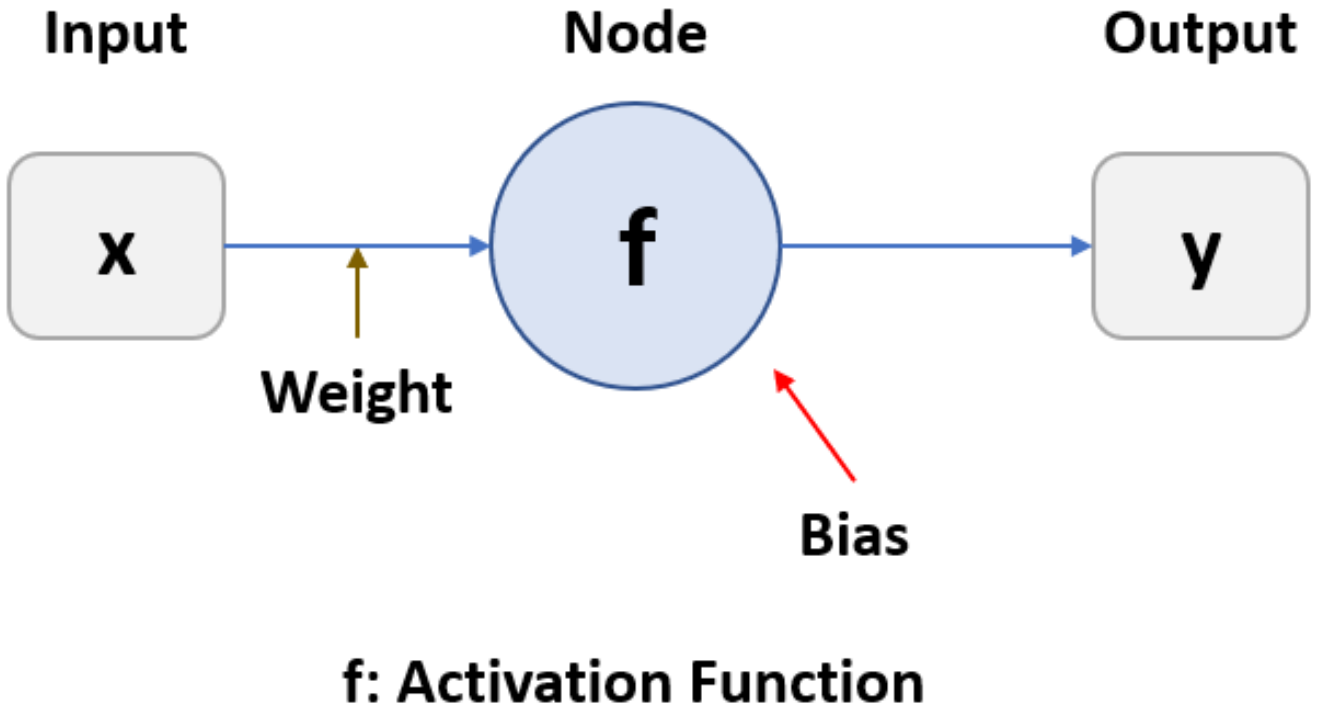

Recommended
Based on what's in this playlist
10,000 Hours (with Justin Bieber)
Dan + Shay, Justin Bieber
Good Things
ADD
2002
Anne-Marie
Speak Your Mind (Deluxe)
ADD
Rewrite The Stars
Zac Efron, Zendaya
Rewrite The Stars
ADD
Lose You To Love Me
Selena Gomez
Rare
ADD
Rewrite The Stars (with James Arthur & Anne-Marie)
James Arthur, Anne-Marie
Rewrite The Stars (with James Arthur & A...
ADD
If I Can't Have You
Shawn Mendes
Shawn Mendes (Deluxe)
ADD
You & Me
James TW
Chapters
ADD
Love Someone
Lukas Graham
3 (The Purple Album)
ADD
Someone You Loved
Lewis Capaldi
Divinely Uninspired To A Hellish Extent
ADD

SONOS

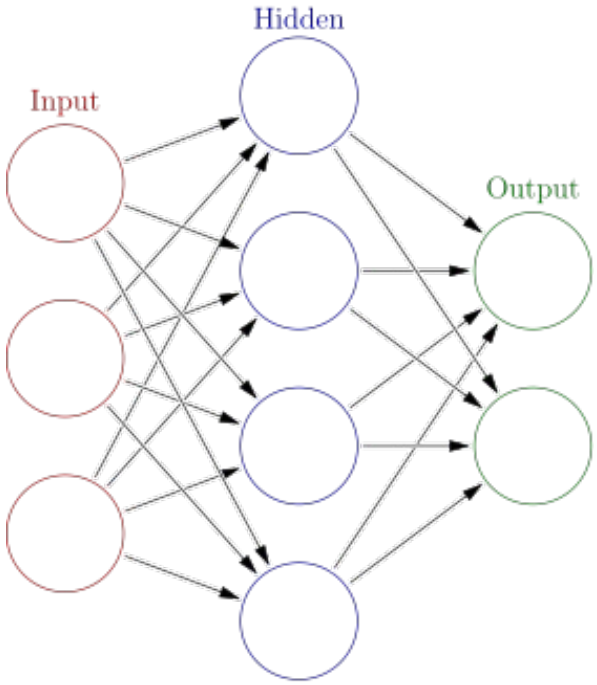
Biological Neural Network



Neuron



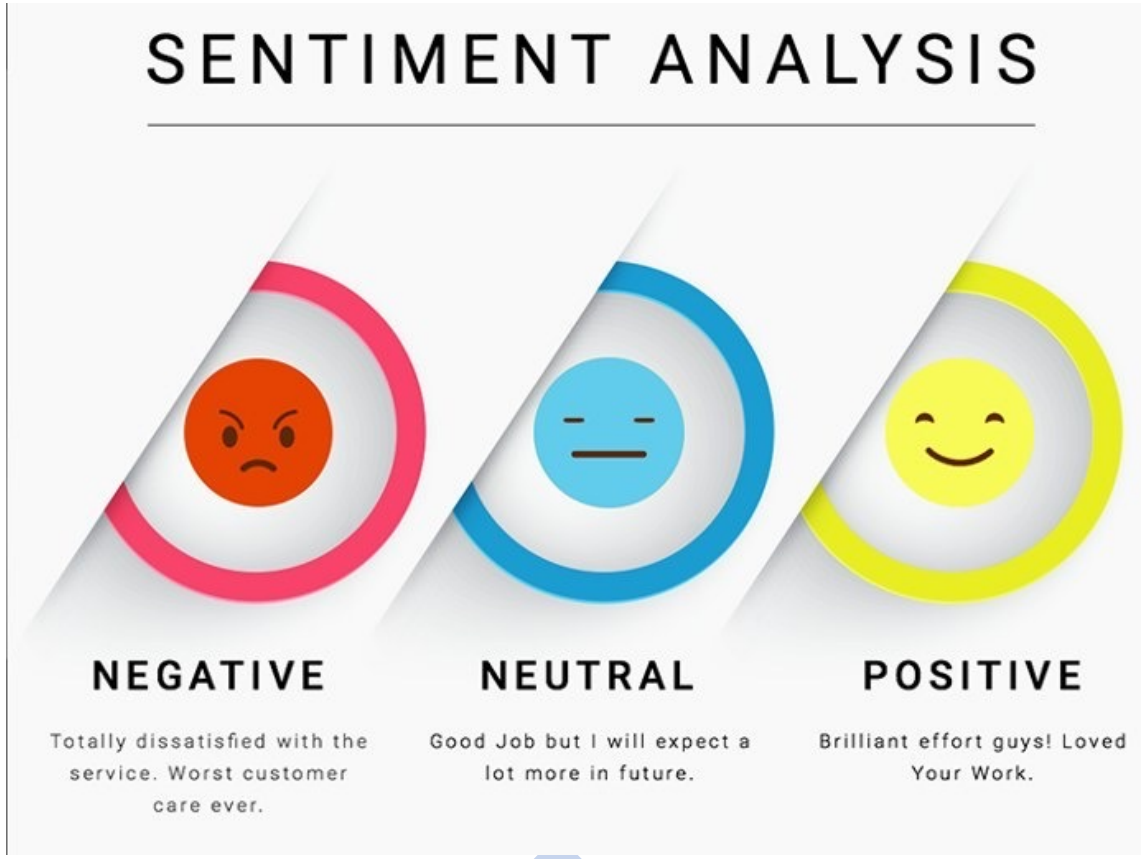
Neural Network



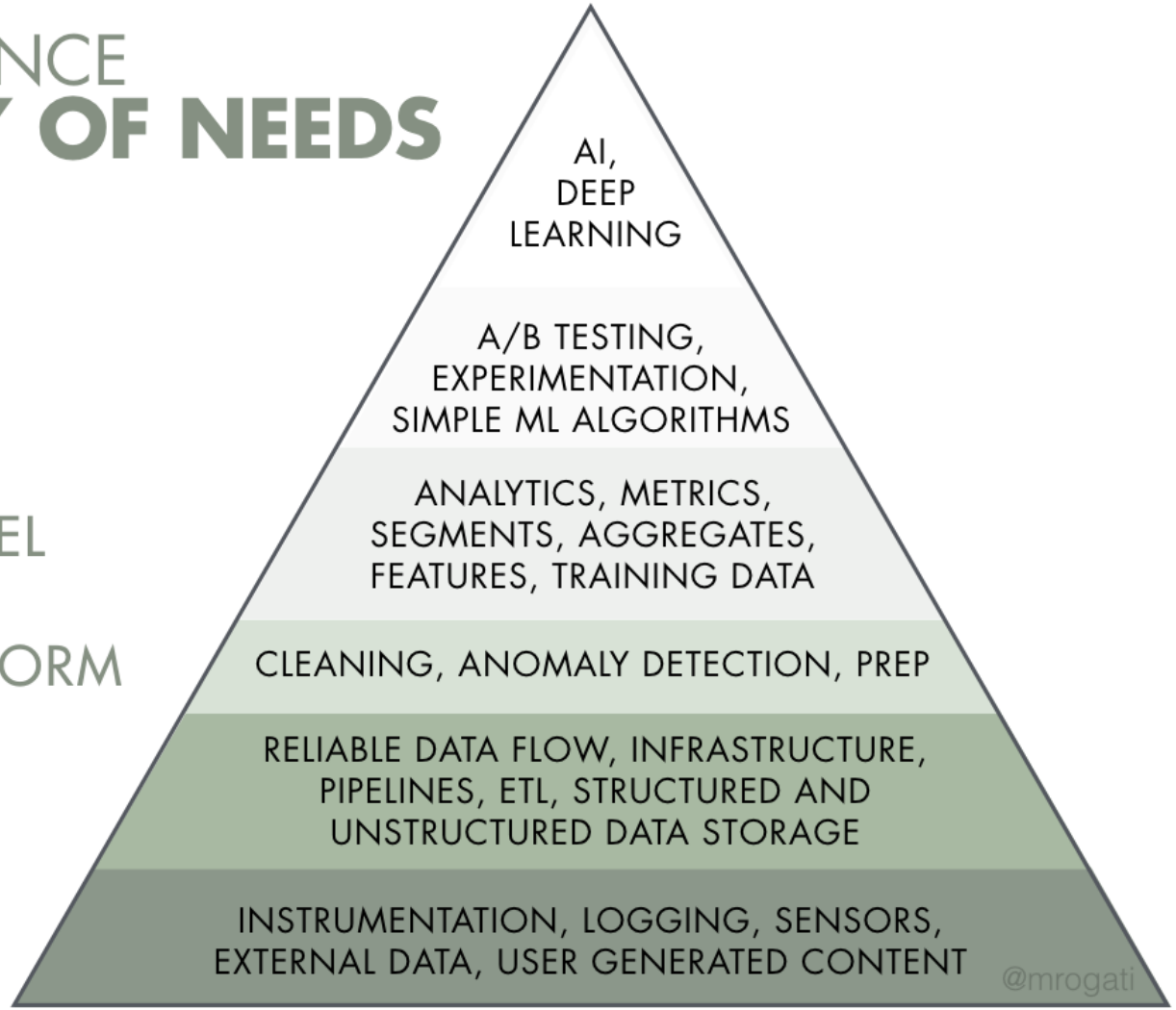
Is this a cat?
Yes or No

Deep learning consists of at least 2 hidden layers.

Machine Learning



THE DATA SCIENCE **HIERARCHY OF NEEDS**



LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT

What Should We Do?

Objectives

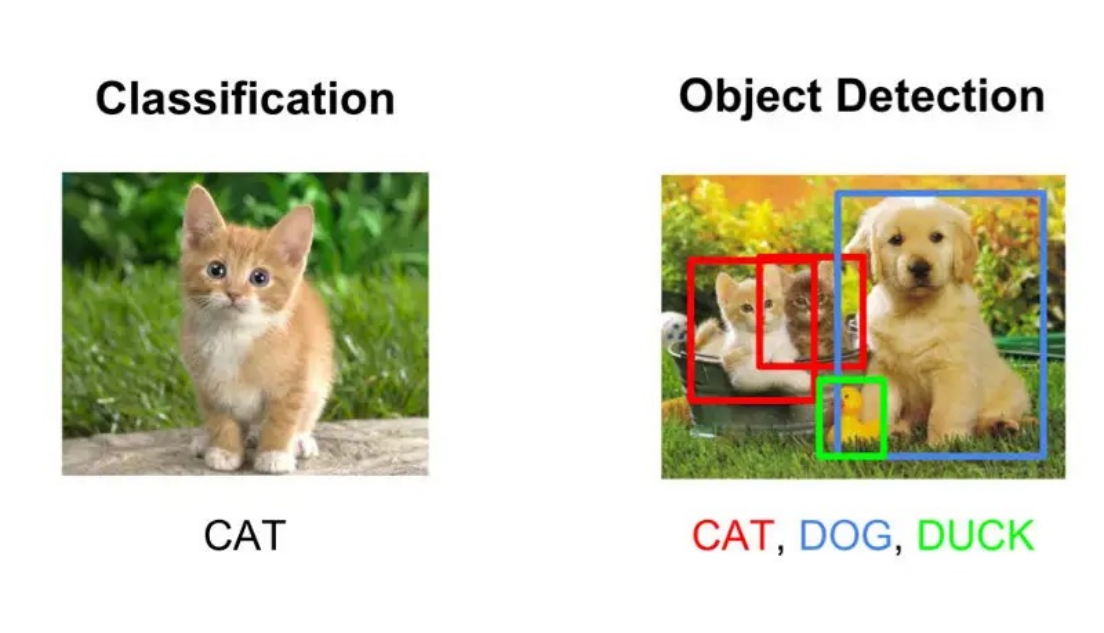
- Data preprocessing
- Fundamental concepts: Activation functions, loss functions, gradient descent, tensor, and normalization.
- Deploy a deep learning model with a dataset.

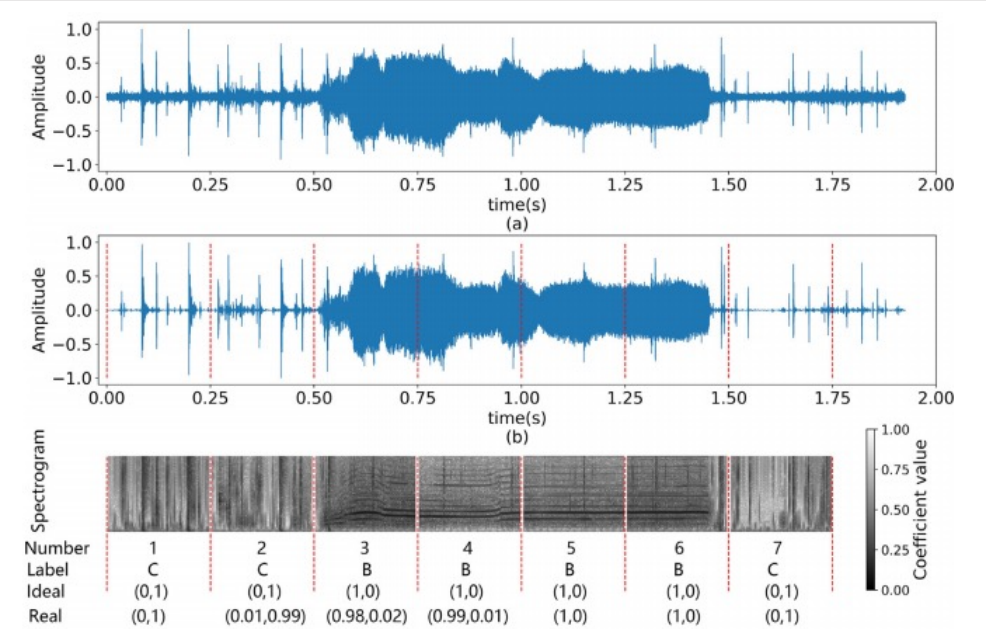
Google Colab

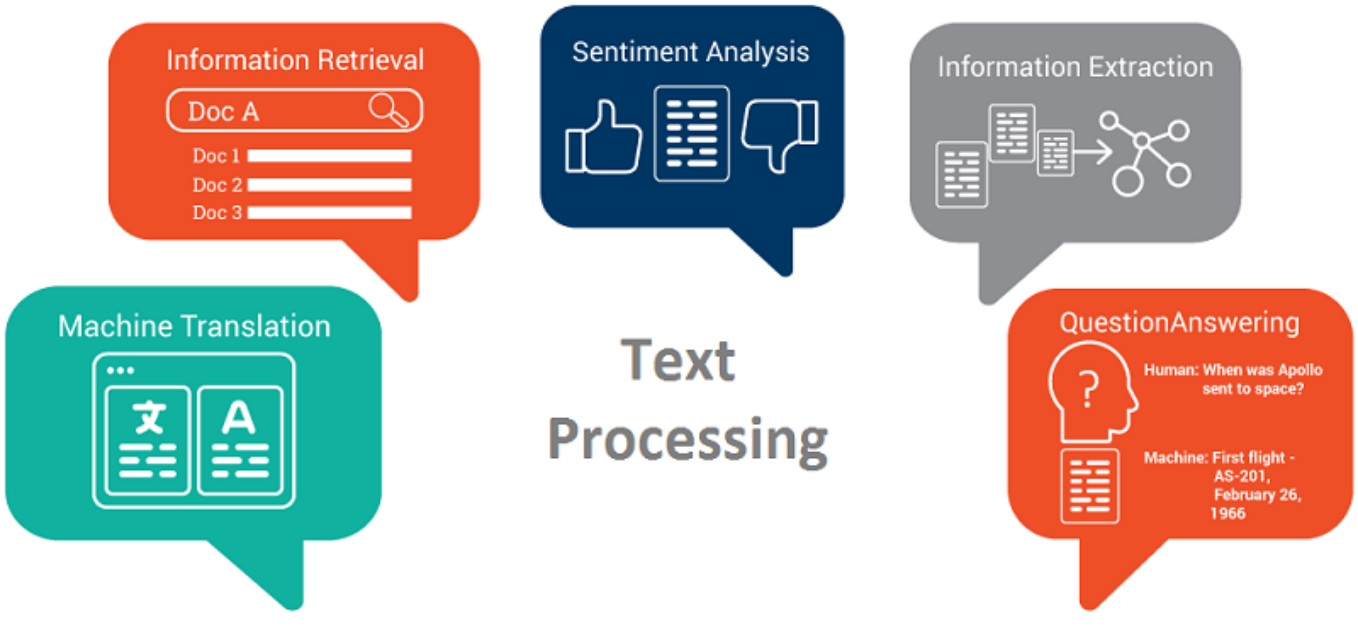
- Google Colab is a cloud computing platform that allows you to run Python code.
- To start Google Colab, go to https://colab.research.google.com/. Once it is set up, go to Google Drive to start Colab.

GitHub

- All code and sources are shared on the GitHub repository (github.com/elibooklover/Deep-Learning-for-Humanists-2023).


Classification
Object Detection
CAT
CAT, DOG, DUCK

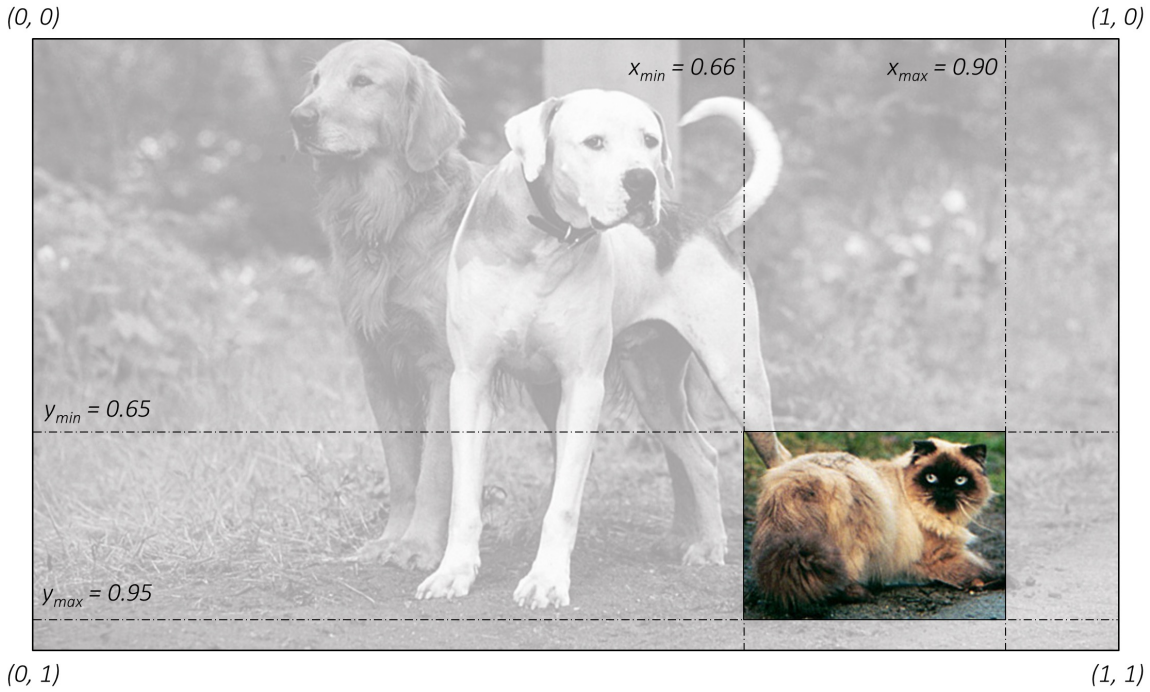

Amplitude
time(s)
(a)
(b)
Spectrogram
Coefficient value
Number 1 2 3 4 5 6 7
Label C C B B B B C
Ideal (0,1) (0,1) (1,0) (1,0) (1,0) (1,0) (0,1)
Real (0,1) (0.01,0.99) (0.98,0.02) (0.99,0.01) (1,0) (1,0) (0,1)


Information Retrieval
Doc A
Doc 1
Doc 2
Doc 3
Sentiment Analysis
Information Extraction
Machine Translation
Text Processing
QuestionAnswering
Human: When was Apollo sent to space?
Machine: First flight - AS-201, February 26, 1966

Computer Vision

Computer Vision Tasks

- Image classification
- Object detection
  - Face recognition
  - Edge detection
- Image restoration
  - Colorization
  - Creating styles
- Video motion analysis



Boundary Coordinates $(x_{min}, y_{min}, x_{max}, y_{max}) = (0.66, 0.65, 0.90, 0.95)$

Classification

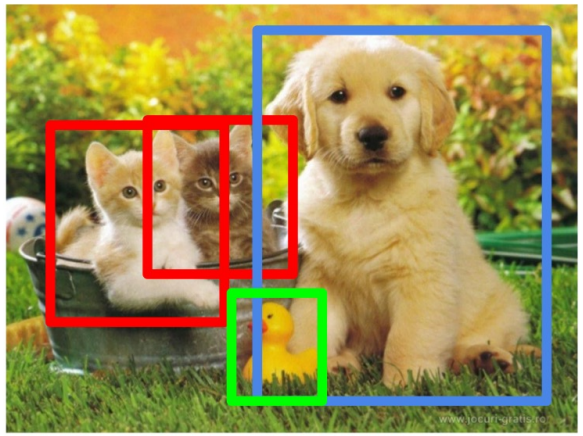
CAT

Classification + Localization

CAT

Single object

Object Detection



CAT, DOG, DUCK

Instance Segmentation

CAT, DOG, DUCK

Multiple objects

Generative Adversarial Networks

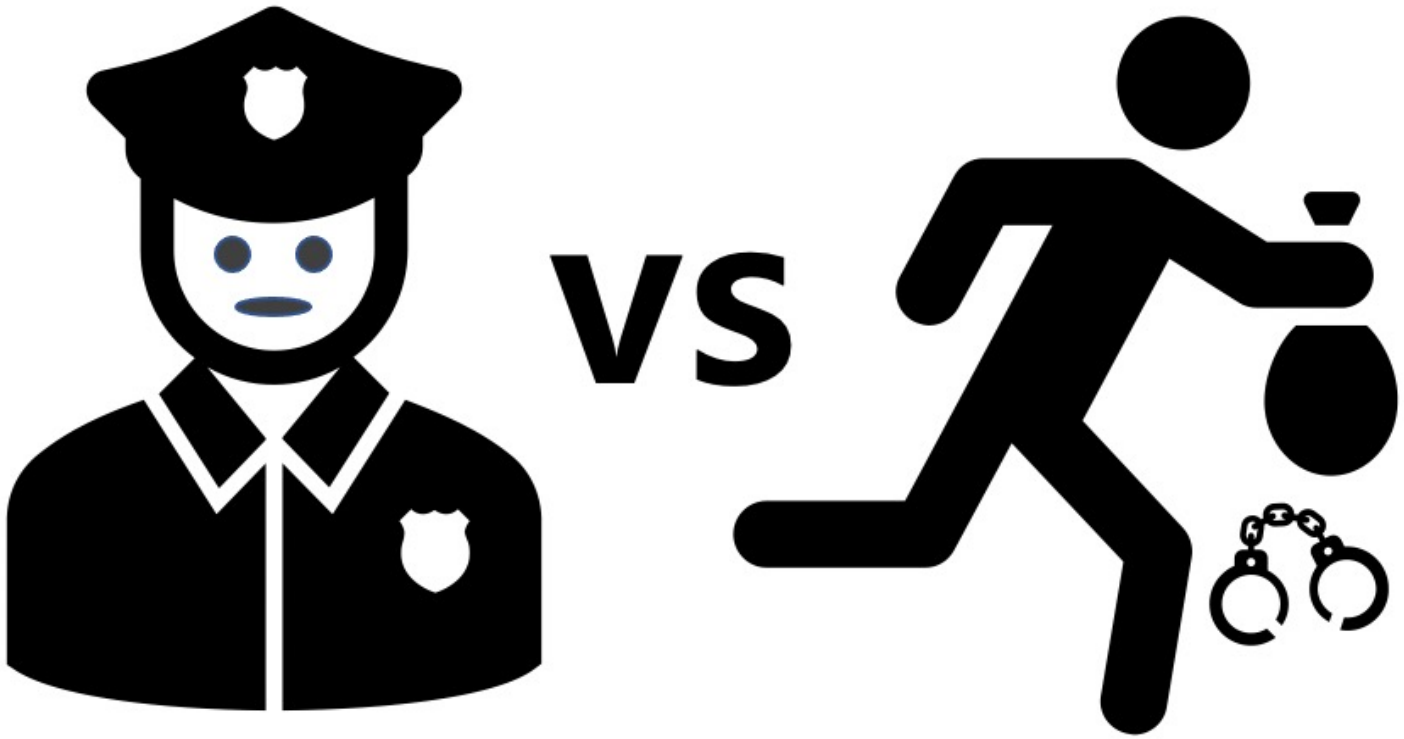
Generative Adversarial Nets

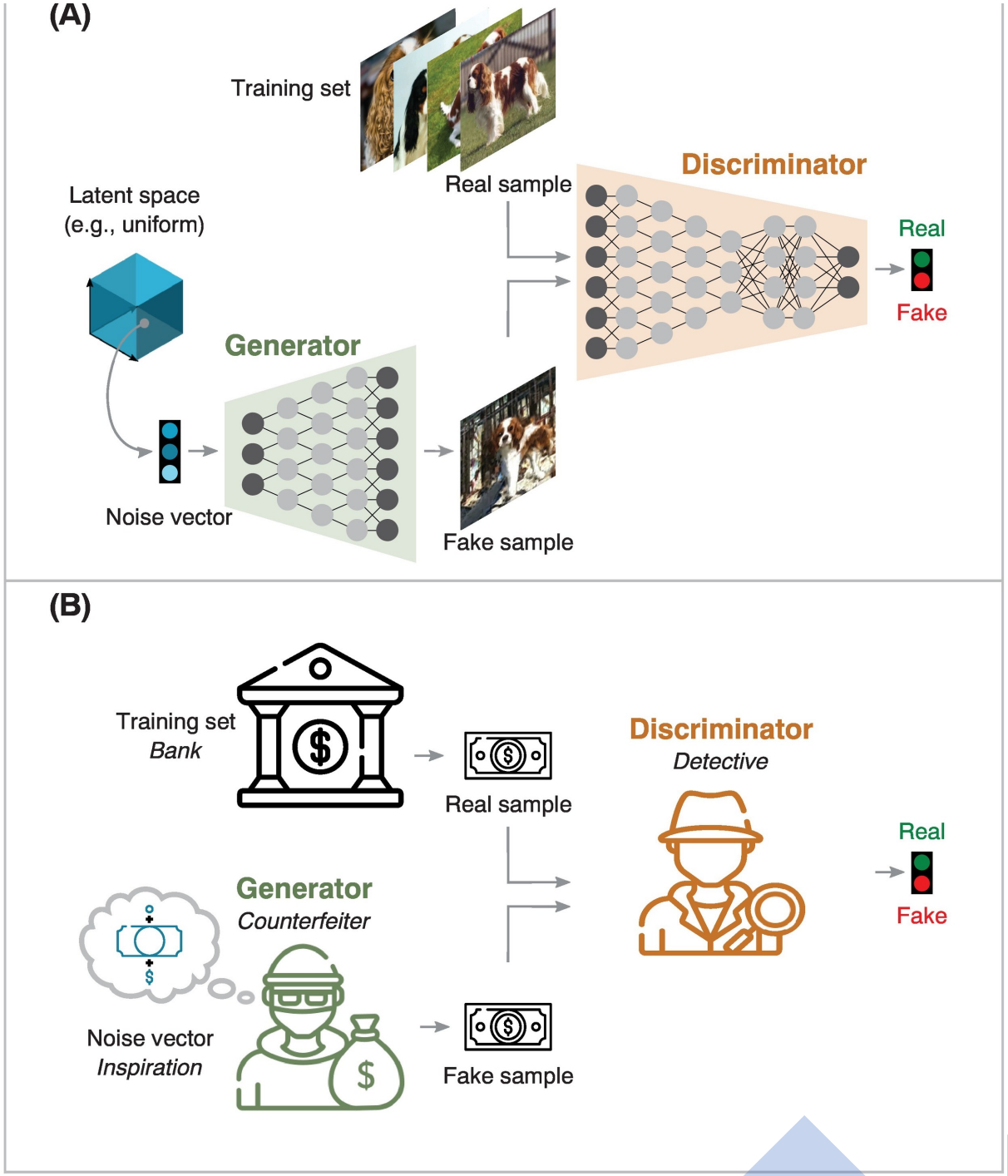
odfellow,* Jean Pouget-Abadie,† Mehdi Mirza, Bing Xu, David Ward
Sherjil Ozair,‡ Aaron Courville, Yoshua Bengio§
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Abstract

propose a new framework for estimating generative models via an adver
process, in which we simultaneously train two models: a generative mode
captures the data distribution, and a discriminative model D that estima
probability that a sample came from the training data rather than G. The tra
procedure for G is to maximize the probability of D making a mistake. T
nework corresponds to a minimax two-player game. In the space of arbitr
ctions G and D, a unique solution exists, with G recovering the training d
ribution and D equal to $\frac{1}{2}$ everywhere. In the case where G and D are defi
nultilayer perceptrons, the entire system can be trained with backpropagati
re is no need for any Markov chains or unrolled approximate inference n
ks during either training or generation of samples. Experiments demonst
potential of the framework through qualitative and quantitative evaluatior
generated samples.

- GAN is a deep learning model to generate fake data (generative modeling).
- Ian Goodfellow et al. introduced GAN for the first time in 2014. (https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf)
- Introduction to GANs at NIPS2016 : youtu.be/9JpdAg6uMXs / www.iangoodfellow.com/slides/2016-12-9-gans.pdf


VS



The image is from
https://doi.org/10.1016/j.tics.2021.06.006

Generative Adversarial Nets

Ian J. Goodfellow,* **Jean Pouget-Abadie,†** **Mehdi Mirza,** **Bing Xu,** **David Warde-Farley,**
Sherjil Ozair,‡ **Aaron Courville,** **Yoshua Bengio**§
Département d'informatique et de recherche opérationnelle
Université de Montréal
Montréal, QC H3C 3J7

Abstract

We propose a new framework for estimating generative models via an adversarial process, in which we simultaneously train two models: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game. In the space of arbitrary functions G and D, a unique solution exists, with G recovering the training data distribution and D equal to $\frac{1}{2}$ everywhere. In the case where G and D are defined by multilayer perceptrons, the entire system can be trained with backpropagation. There is no need for any Markov chains or unrolled approximate inference networks during either training or generation of samples. Experiments demonstrate the potential of the framework through qualitative and quantitative evaluation of the generated samples.

GAN-Based Models

- GAN cannot create high resolution images.
- GAN creates instability of training deep learning models.
- Based on the first GAN model, GAN-based models began to be explored such as Vanilla GAN, cGAN, SRGAN, and DCGAN.

Conditional Adversarial Net

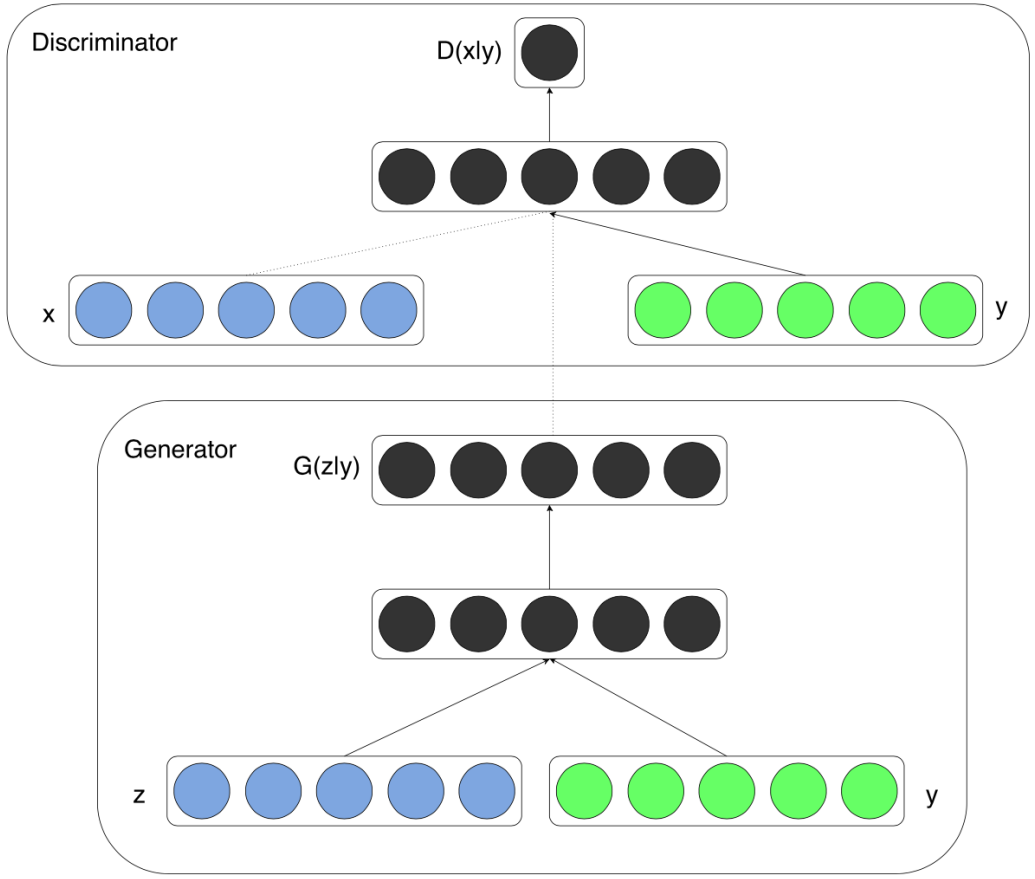
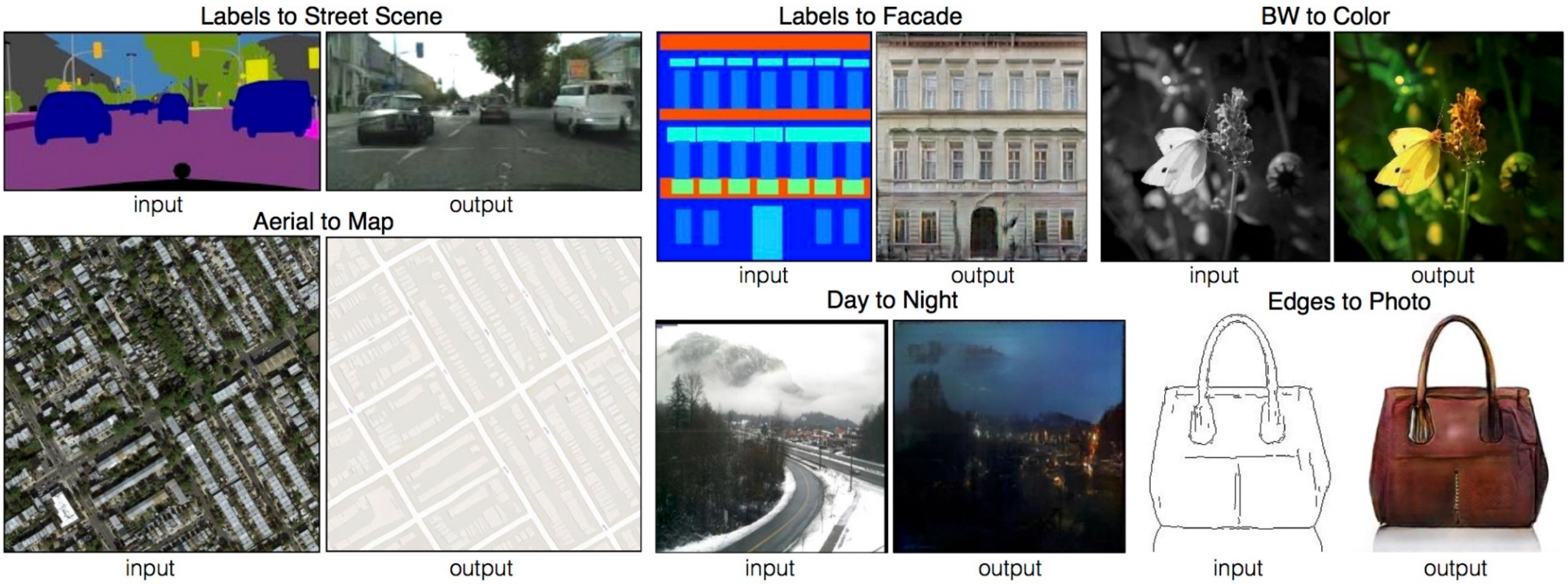


Image: arxiv.org/pdf/1411.1784.pdf

pix2pix



Example results on several image-to-image translation problems. In each case we use the same architecture and objective, simply training on different data.

pix2pix

- “We investigate conditional adversarial networks as a general-purpose solution to image-to-image translation problems. These networks not only learn the mapping from input image to output image, but also learn a loss function to train this mapping. This makes it possible to apply the same generic approach to problems that traditionally would require very different loss formulations. We demonstrate that this approach is effective at synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images, among other tasks” (Isola et al. 1)
- Interactive demo: https://affinelayer.com/pixsrv/

CycleGAN

Converting Monet into Thomas Kinkade

What if **Claude Monet** had lived to see the rise of Americana pastoral kitsch in the style of **Thomas Kinkade**? And what if he resorted to it to support himself in his old age? Using CycleGAN, our great **David Fouhey** finally realized the dream of Claude Monet revisiting his cherished work in light of Thomas Kinkade, the self-stylized painter of light.

Resurrecting Ancient Cities

Jack Clark used our code to convert ancient maps of **Babylon**, **Jerusalem** and **London** into modern Google Maps and satellite views.

Animal Transfiguration

Tatsuya Hatanaka trained our method to translate black bears to pandas. See more examples and download the models at the **website**. **Matt Powell** performed transfiguration between different species of birds

Portrait to Dollface

Mario Klingemann used our code to translate portraits into dollface. See how the characters in Game of Thrones look like in the doll world.

Face ↔ Ramen

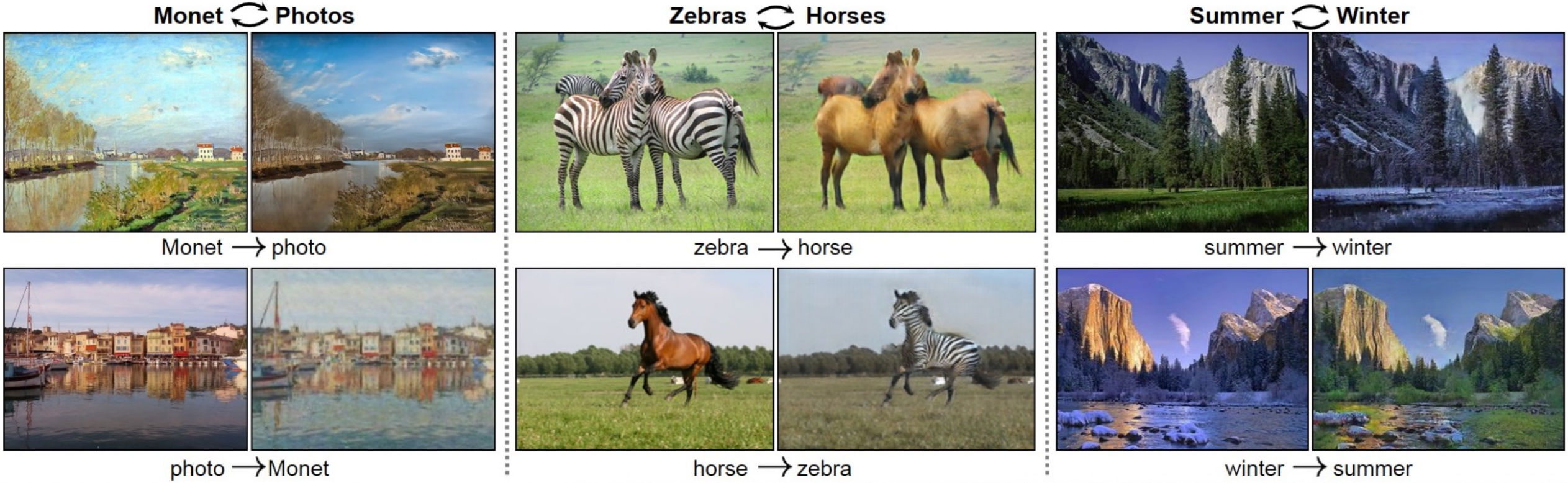
Takuya Kato performed a magical and hilarious Face ↔ Ramen translation with CycleGAN. Check out more results **here**

Colorizing legacy photographs

Mario Klingemann trained our method to turn legacy black and white photos into color versions.

Source: https://junyanz.github.io/CycleGAN/

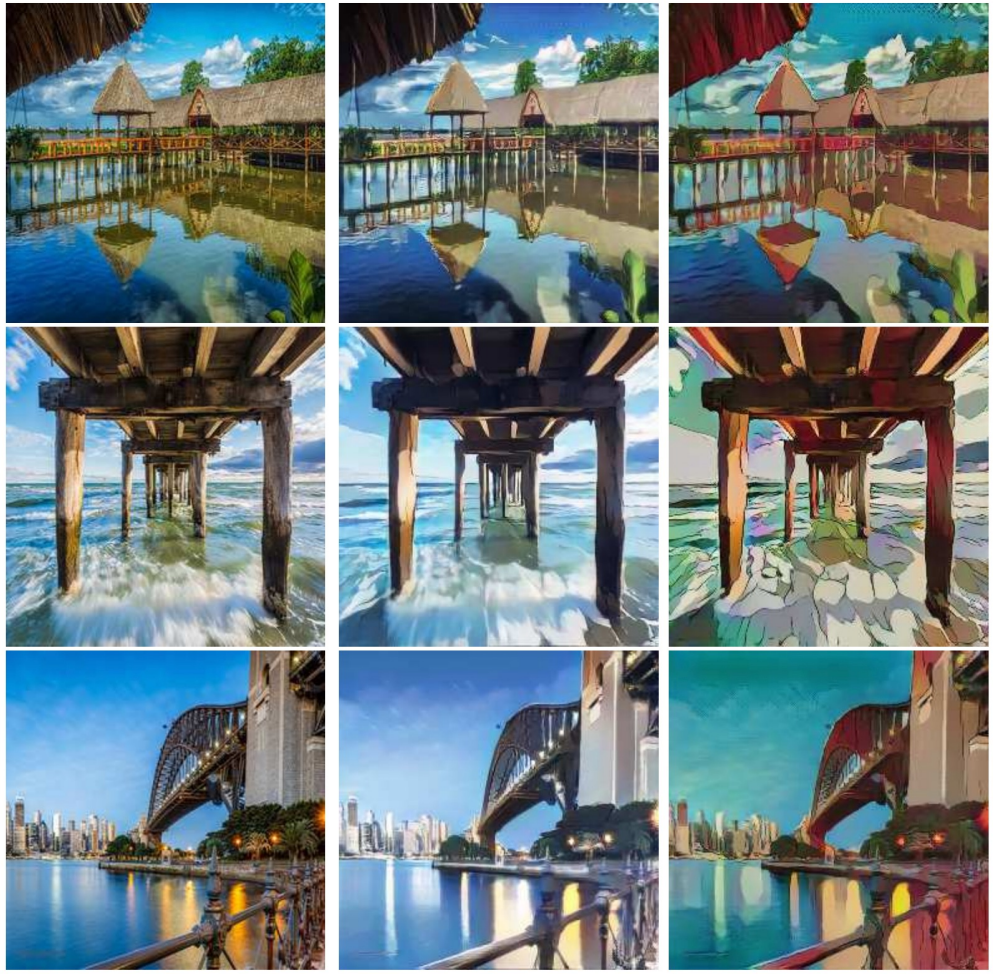
CycleGAN



Source: https://junyanz.github.io/CycleGAN/

Input
Monet
Van Gogh
Cezanne
Ukiyo-e

CartoonGAN

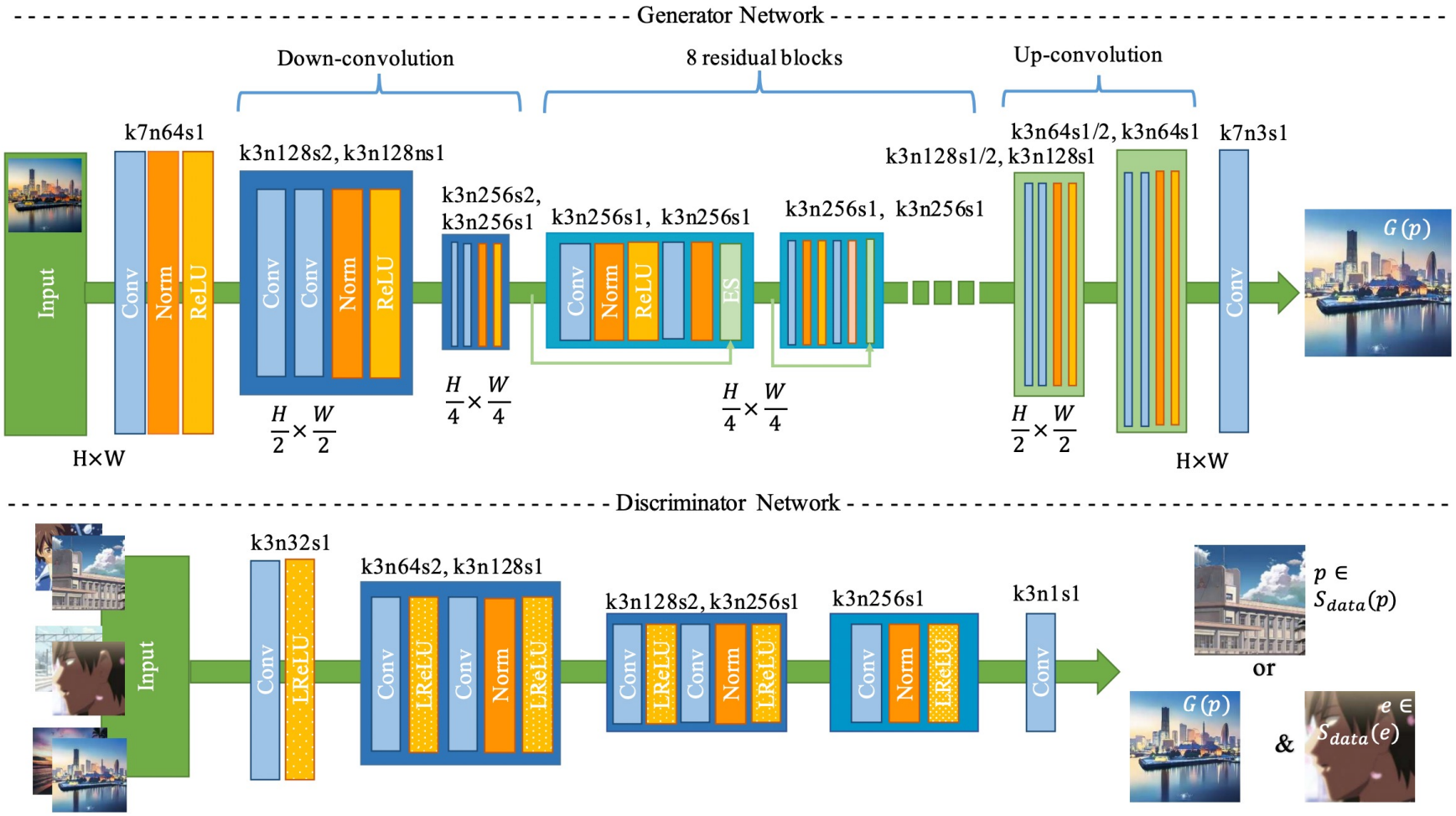


(a) input photo (b) Shinkai style (c) Hayao style

- Generative Adversarial Networks for Photo Cartoonization (CVPR 2018)
- A Novel Lightweight GAN for Photo Animation (ISICA 2019)

Image: https://openaccess.thecvf.com/content_cvpr_2018/papers/Chen_CartoonGAN_Generative_Adversarial_CVPR_2018_paper.pdf

CartoonGAN



encoder-decoder

PatchGAN

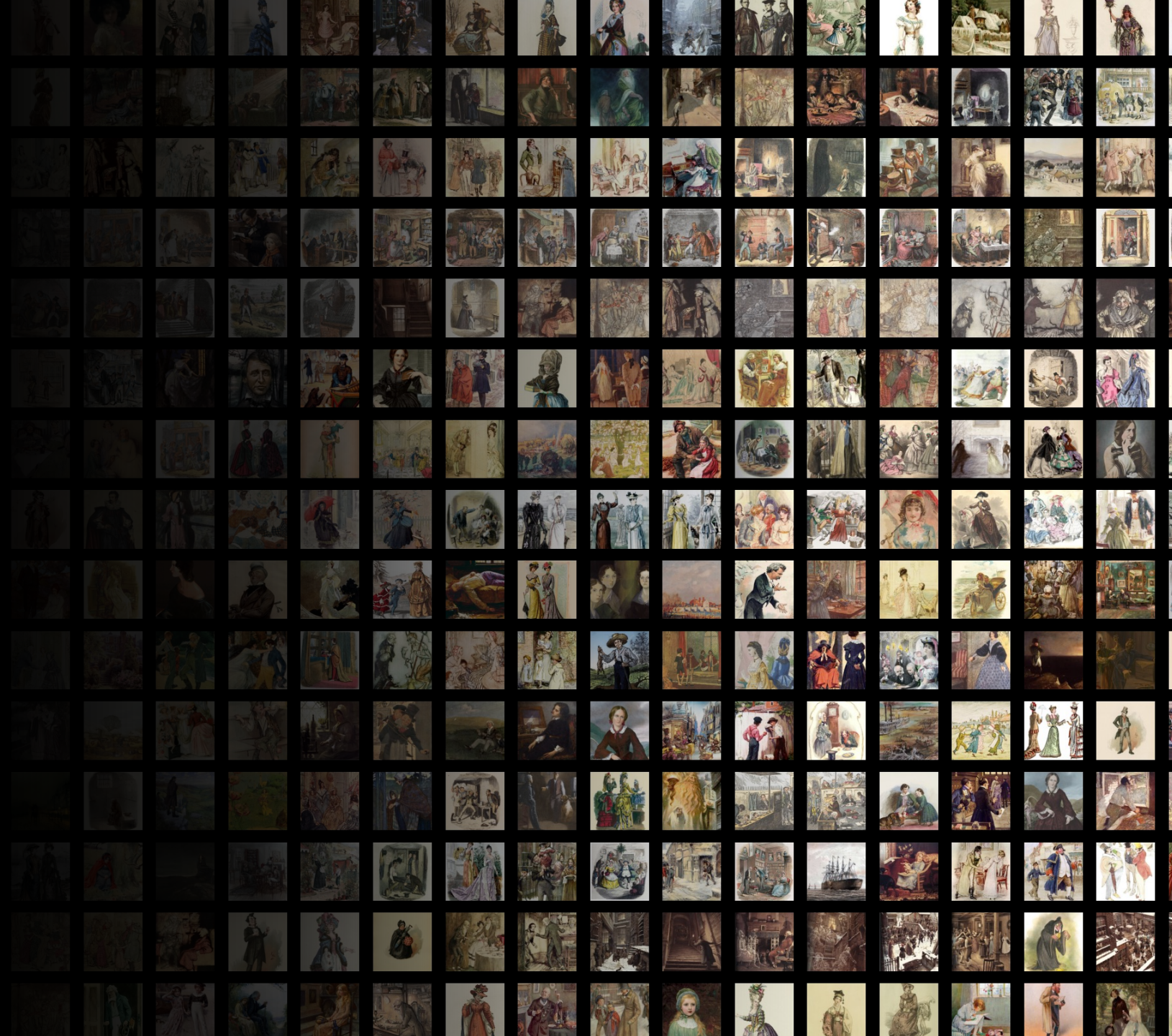
Figure 2. Architecture of the generator and discriminator networks in the proposed CartoonGAN, in which k is the kernel size, n is the number of feature maps and s is the stride in each convolutional layer, 'norm' indicates a normalization layer and 'ES' indicates elementwise sum.

Image: https://openaccess.thecvf.com/content_cvpr_2018/papers/Chen_CartoonGAN_Generative_Adversarial_CVPR_2018_paper.pdf

Victorian400

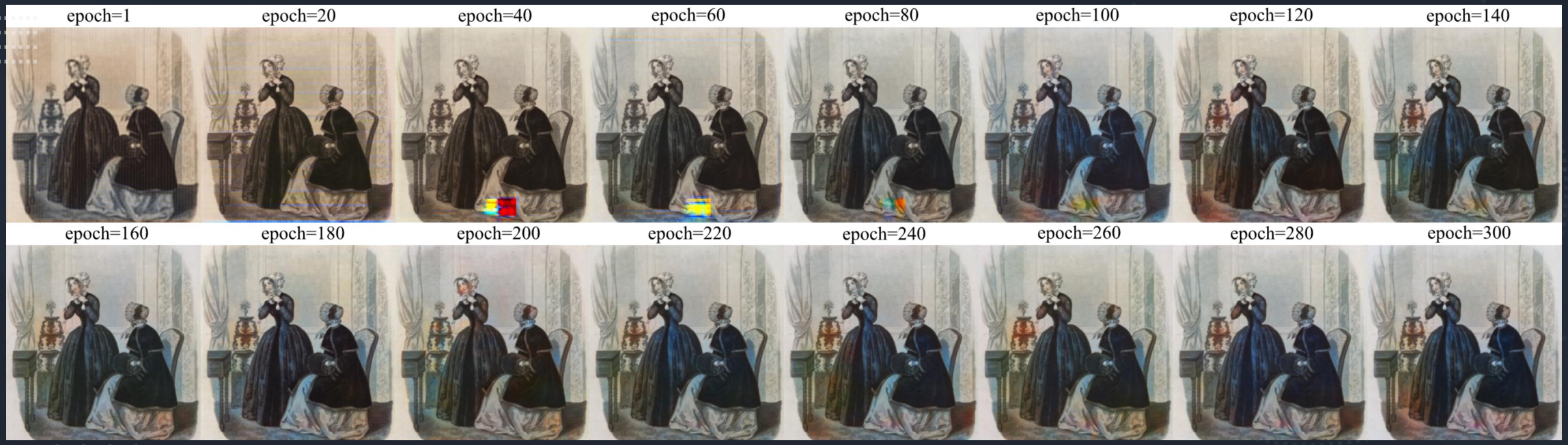
Victorian400

- Victorian400 is a nineteenth century illustration dataset consisting of 400 colorful images.
- Victorian400 helps to run code easily, test and develop deep learning models.
- Victorian400 shows possibilities for improving students' ability to better understand the intersections of text and image.
- I selected and curated images for the Victorian400 dataset in order to improve the accuracy of the colorization results.



Victorian400

- I tested the pix2pix model, which is a conditional generative adversarial network (cGAN) model, with the Victorian400 dataset to verify the possibilities of colorizing black-and-white illustrations from the nineteenth century.
- Based on the test result, I chose 200 epochs to predict colors for black-and-white illustrations.



Input Output Target

Input Output Target

TensorFlow

PyTorch

Machine Learning Frameworks

Any Questions?