Lecture 1

- Generation (Graphics) vs Inference (vision)
- Statistical generative models = probability distribution p(x)
  - Data : Samples (e.g. : images of bedrooms)
  - Prior knowledge: parametric Gaussian, loss function, etc.
- It is generative as you can sample from p(x) --> generate new samples
- Discriminative models → generic ML which is more or like classification or regression (Conditional distribution)
- Model the combinations of input, output basis the probabilities in generative models
- *Generative can handle missing data?? How?*
- Conditional generative models → given a condition, generate the corresponding sample
- Generative is Low dimension input to high dimension output and discriminative is vice versa
- GAN/VAE → Learn a mapping from a low dimensional space to the data manifold(high dimensional space) using neural networks
- *Deep generative models (VAEs, GANs) may also serve for data compression For a given X, we can look up it's latent variable (or code) z? How??*
- Gaussian distributions or non-gaussian?
- AutoRegressive models (used for sequential data often) - with the use of existing audio data as input, new samples are generated → WaveNet (can be used for Text to Speech)
- Image Super-Resolution (enhancing quality of video or images → by generating subsamples in high resolution using generative models)
- Audio Super-resolution similarly to enhance videos
- Machine Translation → language translation
- Image Translation
- Learning Disentangled Representations
- Image Compression
- Unsupervised is difficult in generative compared to discriminative models
- *Benjo Deep learning book*

Lecture 2

- Bernoulli (coin flip) vs categorical distribution(m sided dice)
- Joint Distribution → example of MNIST dataset
- Marginal distributions and independence
  - Independence assumption is very strong and model is not likely to be useful
  - For a pixel image, the possible states are 2^n but assuming independence we can reduce the parameters needed to be n
- Conditional independence

    - Example: a 3rd variable can make 2 variables conditional independent, which could be correlated but not independent. Like height and vocab of children, with age being the 3rd variable
- Probability rules
    - Chain Rule
    - Bayes Rule
    - Markov assumption → used for time series → to find the next step, you only need to find the present value
        - p(x1, x2, , xn) = p(x1)p(x2/x1)(x3/x2)...p(xn/xn-1)
        - The parameters needed are 2n-1
    - Bayesion Networks → directed acyclic graph (Joint Distribution)
        - ***Stochastic ?***
        - ***Causal Inference?***
        - Without conditional independence, the node could be correlated with each other. So its necessary to condition on the parent nodes
    - Neural networks can be used within bayesian networks for parameterization

Lecture 3

- Learning generative model
    - Generation
    - Density estimation
    - Unsupervised representation learning
- Bayesian networks vs neural models
    - Neural networks can be replaced as subelements in bayesian networks
- Pixel CNN's → looks at surrounding pixels to generate the next
- Masking allows autoencoders to be auto-regressive models
- How did voice generation happen? Wavenet read about it

Adversial training -
https://towardsdatascience.com/adversarial-examples-in-deep-learning-be0b08a94953
GAN -
https://towardsdatascience.com/generative-adversarial-networks-gans-2231c5943b11

Binomial distributions
Gaussian distributions
Lagrange;s
Law of large numbers
Nash equilibrium
https://towardsdatascience.com/adversarial-examples-in-deep-learning-be0b08a94953

https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html

https://github.com/eriklindernoren/PyTorch-GAN/blob/master/implementations/wgan/wgan.py
https://towardsdatascience.com/build-a-super-simple-gan-in-pytorch-54ba349920e4

https://towardsdatascience.com/building-a-gan-with-pytorch-237b4b07ca9a

DistilBERT introduction:
https://jalammar.github.io/a-visual-guide-to-using-bert-for-the-first-time/

PCA:
https://towardsdatascience.com/a-one-stop-shop-for-principal-component-analysis-5582fb7e0a9c

VAE:
https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73