Leveraging Generative AI for Advanced Scientific Research
Breakthroughs and Barriers

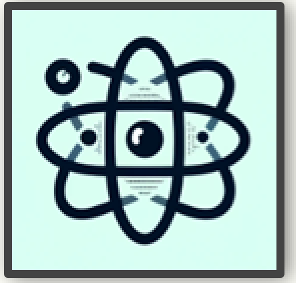
Md Mushfiqur Rahman

PhD Student
Department of Computer Science
George Mason University
mrahma45@gmu.edu

About Me

Making generative machine learning usable for real world applications

Generative NLP in Healthcare



Generative AI in Research

Goal of today’s talk

Is the current state of Machine Learning ready for scientific research?

Overview

- Intro to Machine Learning
- Generative ML
- Application in Research: Concerns and Solutions
- Case Study 1: Large Language Models
- Case Study 2: Image Generation Models
- Conclusion

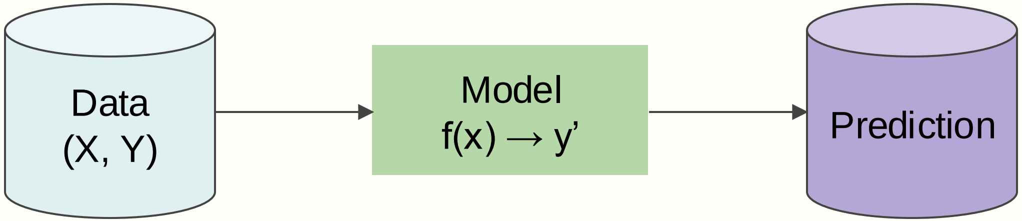
Current State of Machine Learning

- Modern ML models are end-user ready
- Reduces need for human intervention
- Often better than empirical formulae
- At times, even better than humans

A very brief introduction to Machine Learning

Algorithm “learns” patterns from data

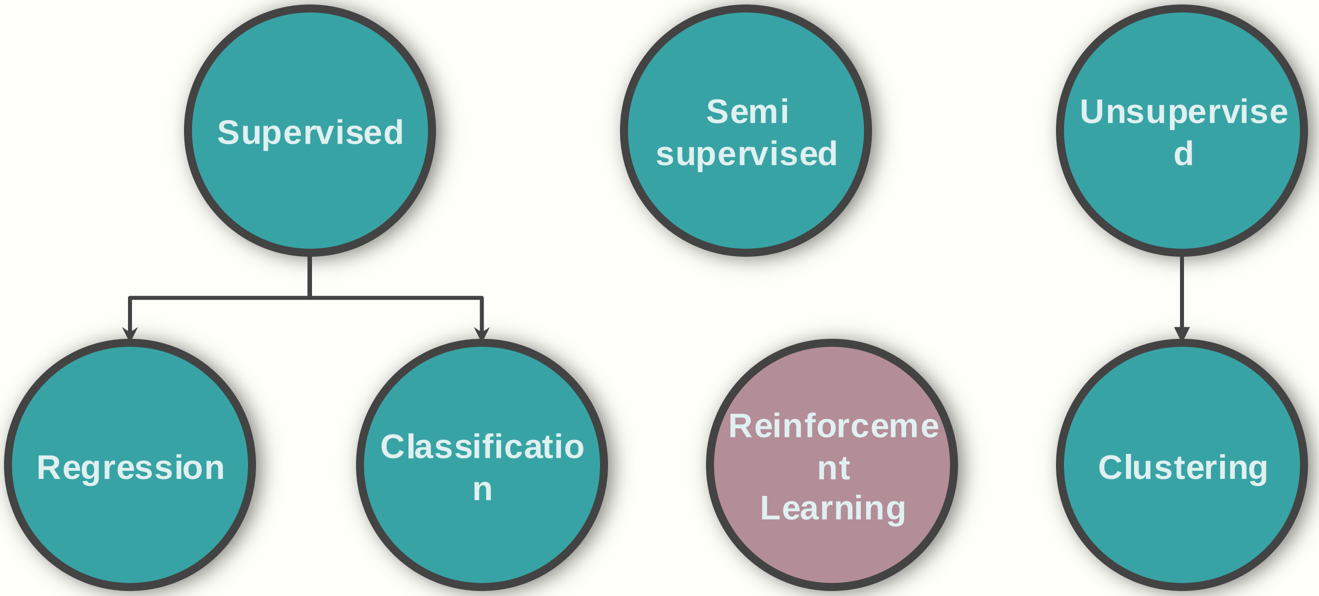
In reality: tries to satisfy objective function by minimizing loss



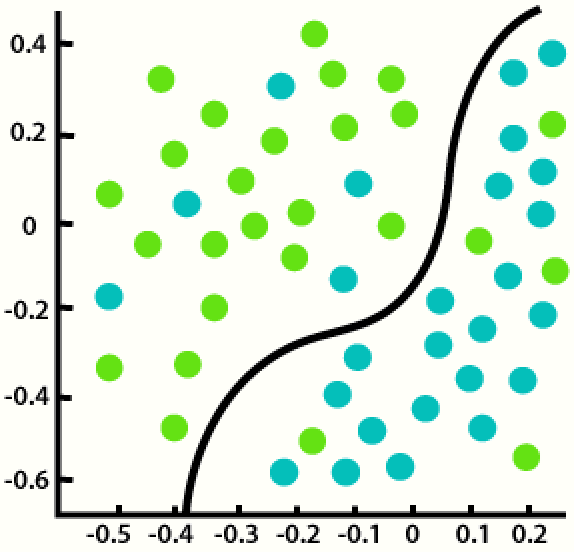
What is a valid input?

- Distinct feature sets
- Images
- Texts
- Almost anything

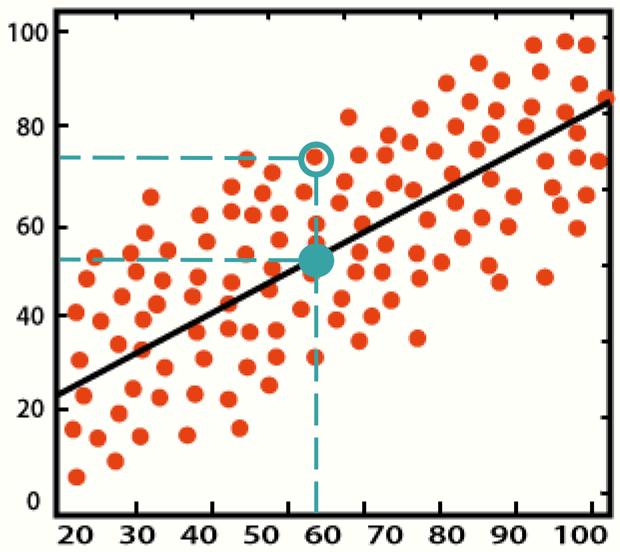
Common Machine Learning Categories



Basic Regression and Classification Models

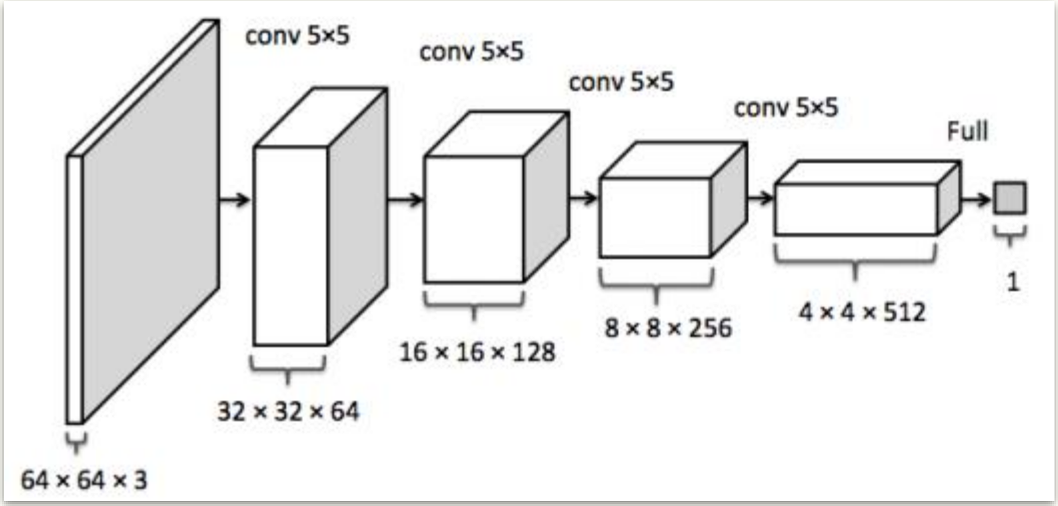


Classification

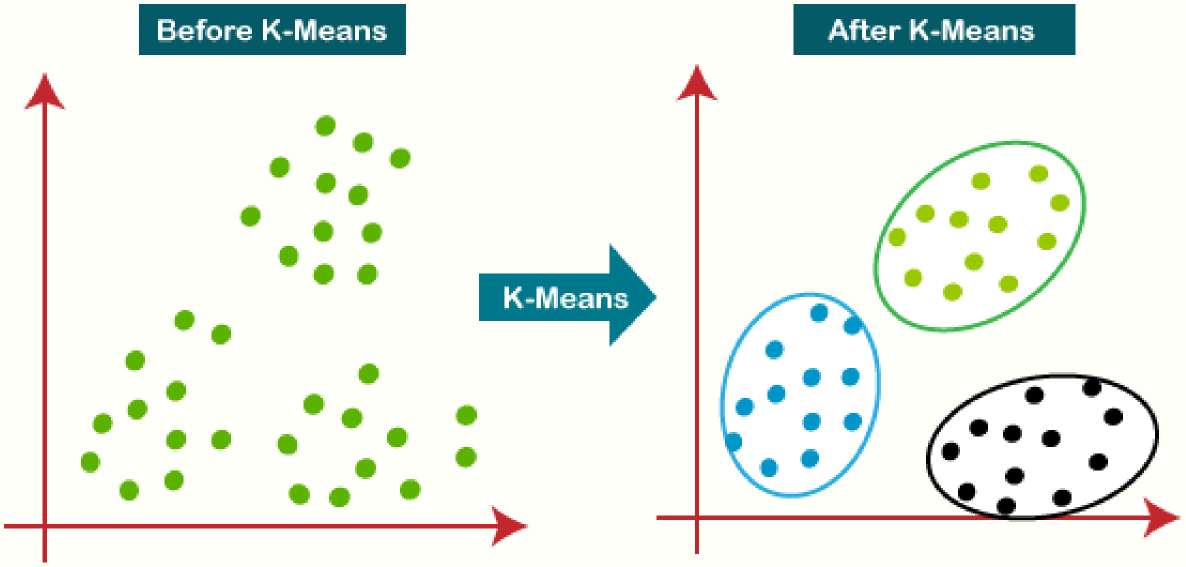


Regression

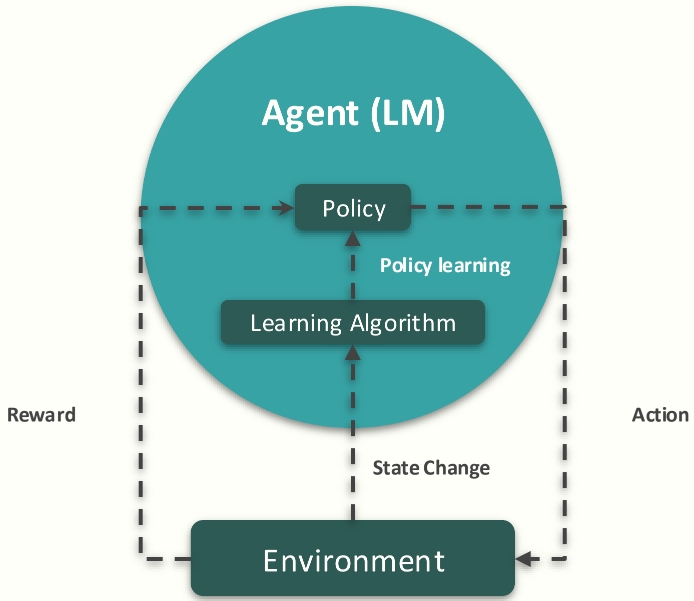
Neural Networks



Clustering Algorithm



Reinforcement Learning

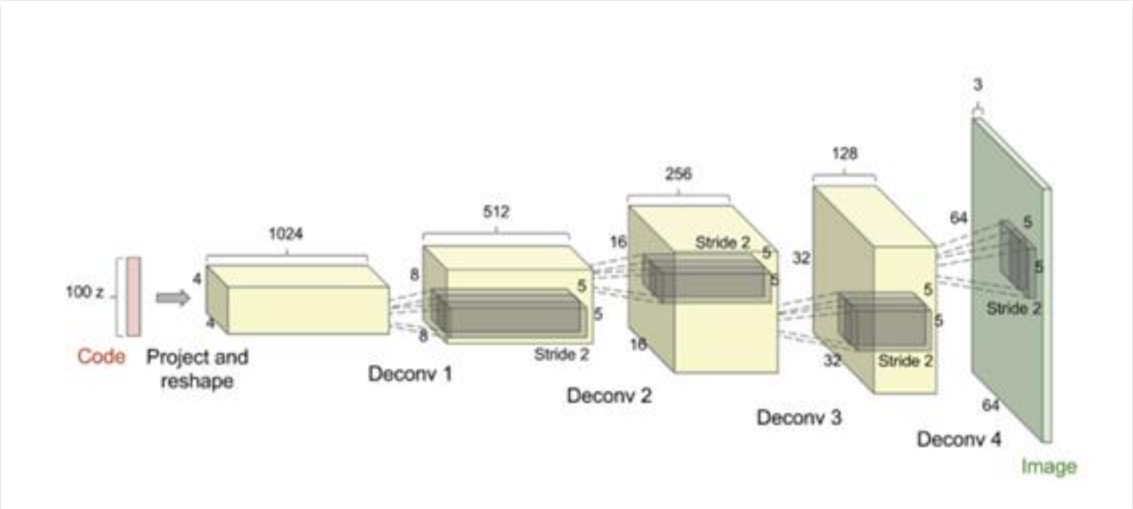


Generative Models

Goal: Generate images or texts

- Synthetic data generation
- Missing data fill-up
- Image segmentation tasks
- Text generation models
- Style transfer tasks

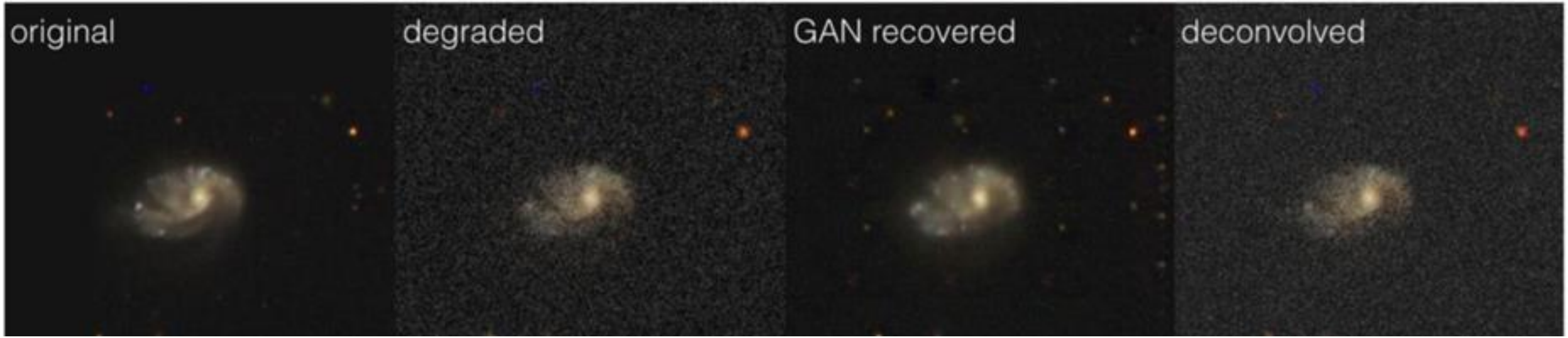
Generative Machine Learning



Much trickier than label prediction

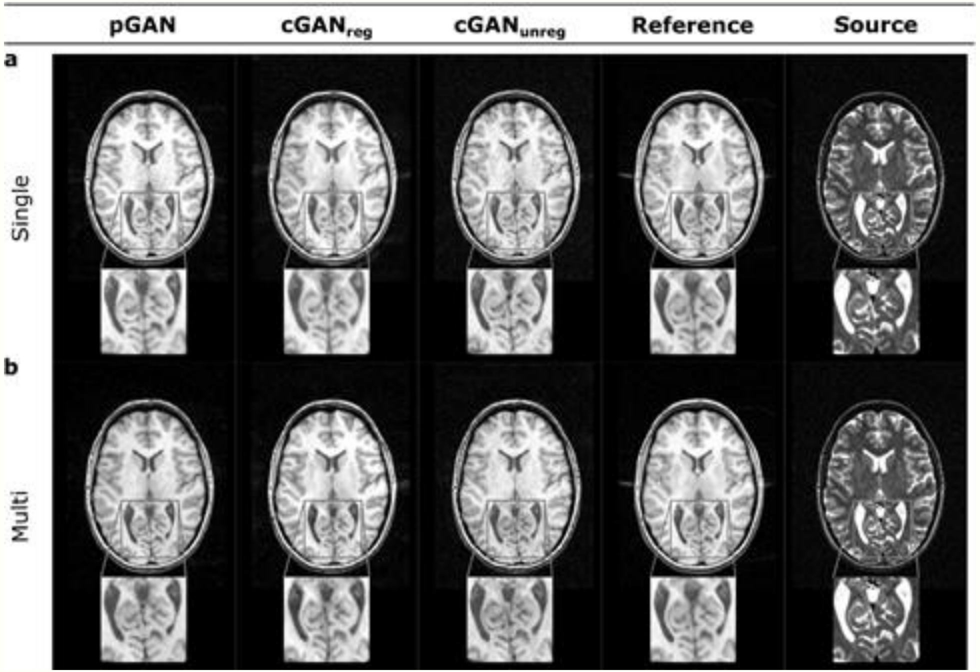
Applications in Scientific Research

Astronomical Image Enhancement



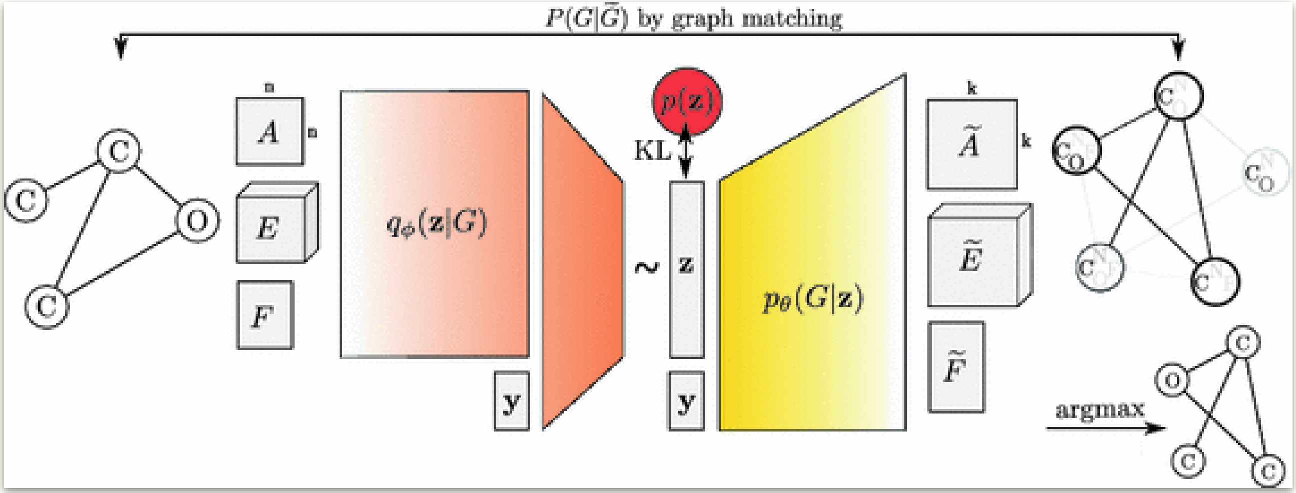
Applications in Scientific Research

Medical Image Synthesis



Applications in Scientific Research

Molecular or Chemical Structure Prediction



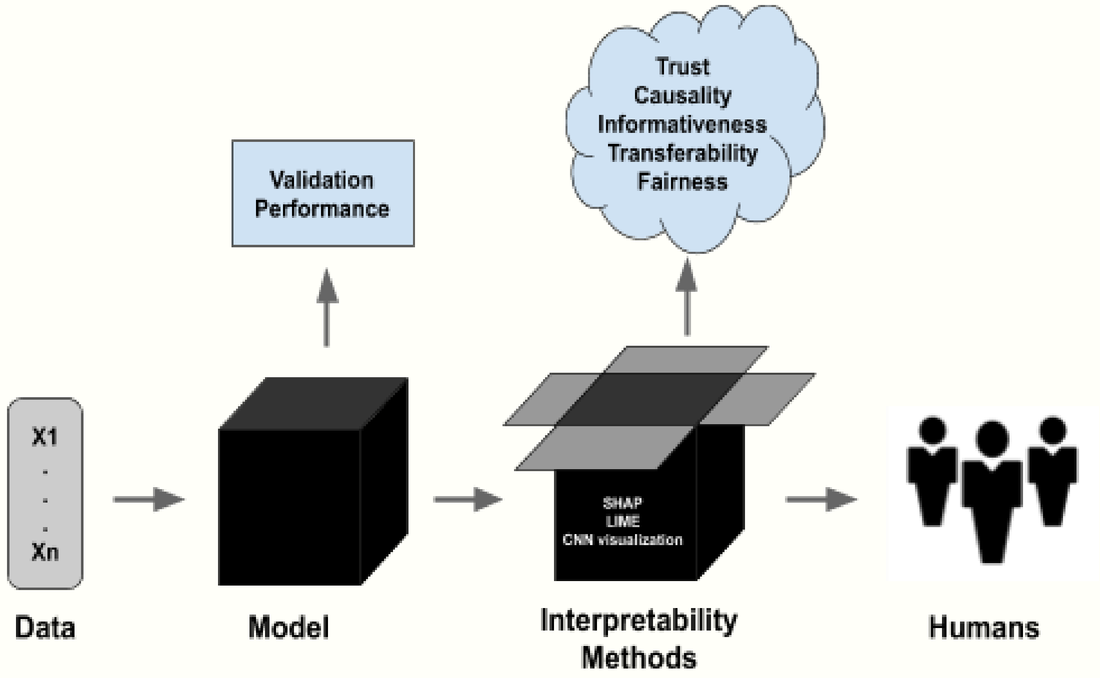
Common Concerns in Scientific Deployment

- Fairness and bias issue
  - Lack of fairness in data reflected on outcome
  - Data collection disparity results in unfair outcomes

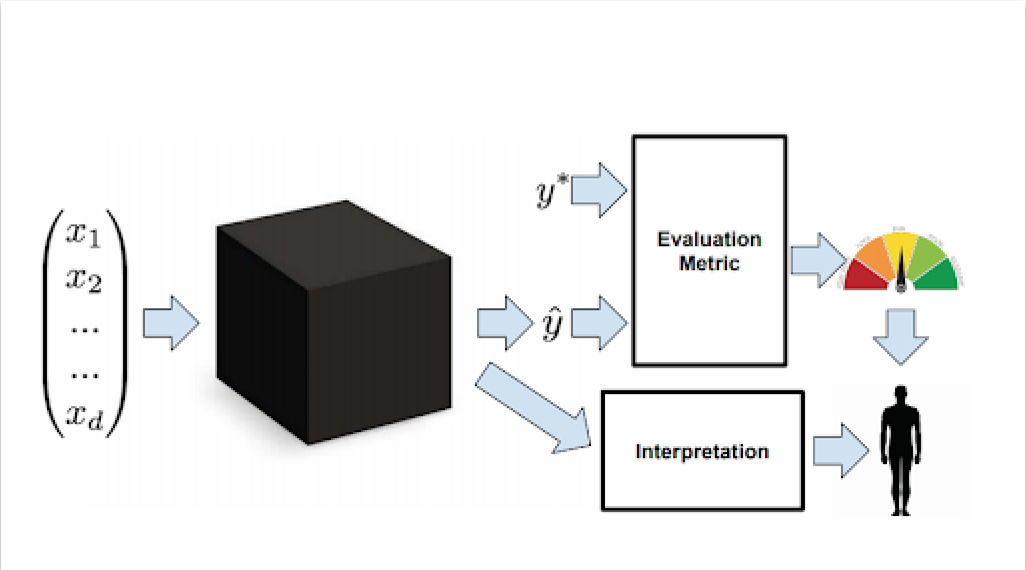
- Models are black-boxes
  - Reasoning behind each decision
  - Issue in evaluating minor components

- Models are unreliable
  - Non-deterministic
  - Issue with reproducibility
  - Heavily reliant on sample data

Model Interpretability



Model Interpretability (contd.)



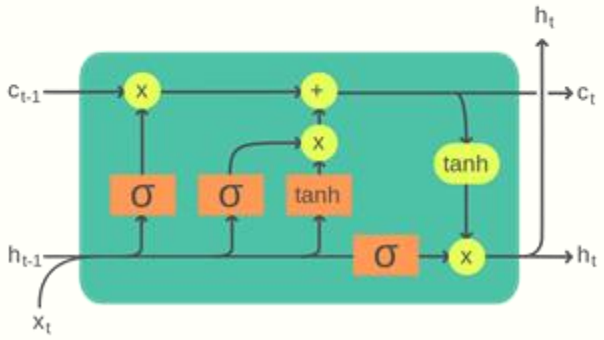
Case Study: A Journey from LSTM to Large Language Models

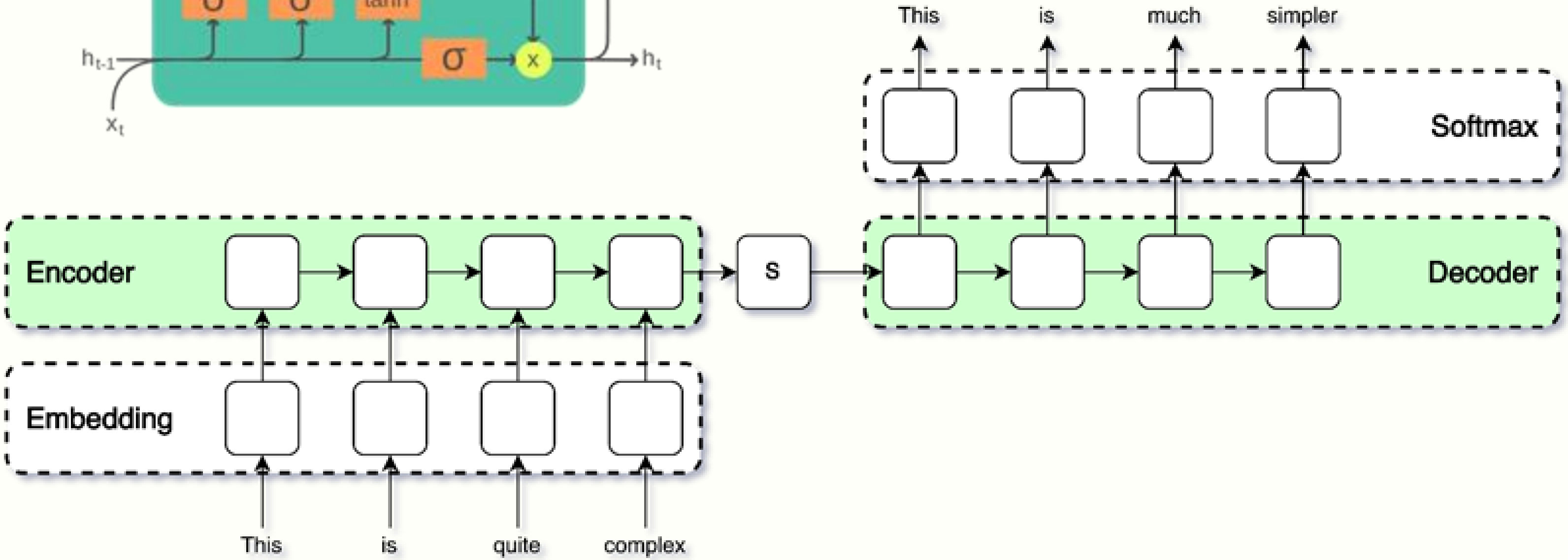
Text Summarization

Machine Translation

Dialogue Generation

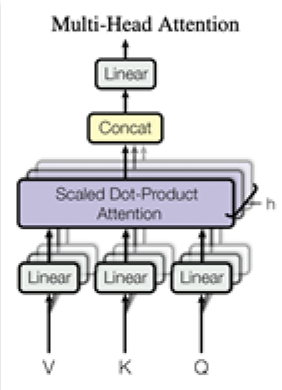
RNN and LSTM

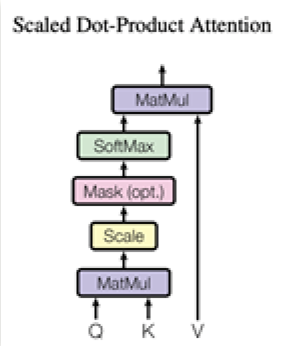


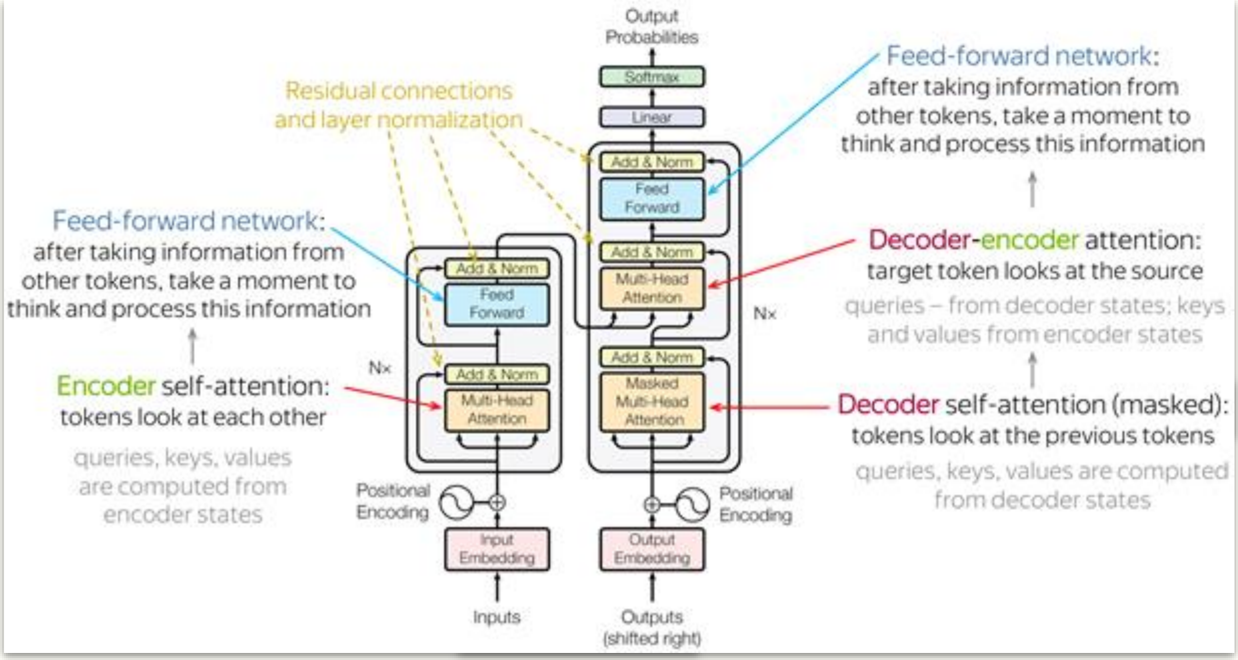


Satskever et al (2014) "Sequence to Sequence Learning with Neural Networks"
Hochreiter et al (1997) "Long Short-Term Memory"

Transformers







$$SelfAttention(\mathbf{Q},\mathbf{K},\mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V}$$

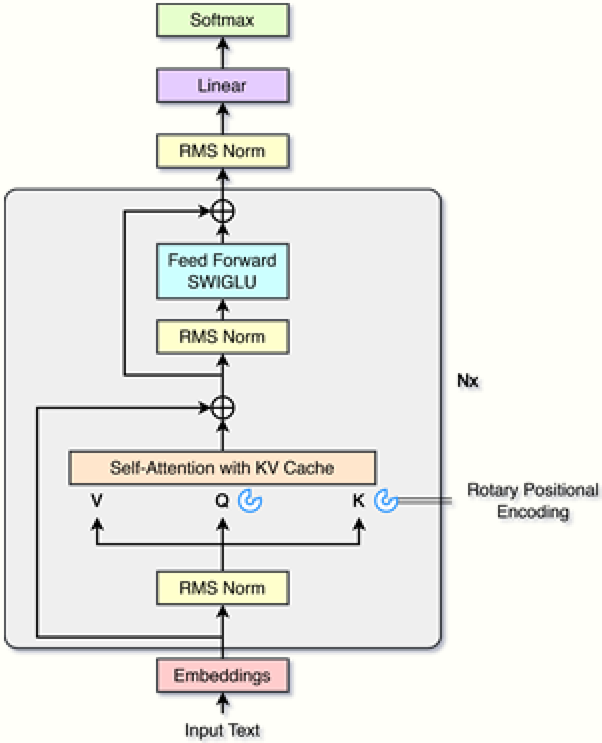
Vaswani et al (2017) “Attention is all you need”
Figure borrowed from a blog by Vaswani

LMs and LLMs with Transformers

- Encoder-based models
  - Generate a rich, contextual representation
  - BERT, RoBERTa

- Decoder-based models
  - Predict the next token
  - Often autoregressive
  - GPT Series, Llama series

- Encoder-decoder models
  - Text-to-text transformers
  - T5, TransforerXL



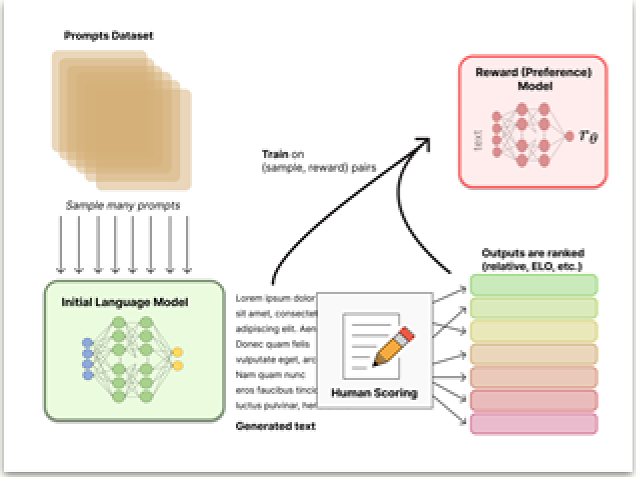
Touvron et al (2023) “LLaMA: Open and Efficient Foundation Language Models”

RL and RLHF

Pre-train LM

Train reward model

Finetune LM with RL



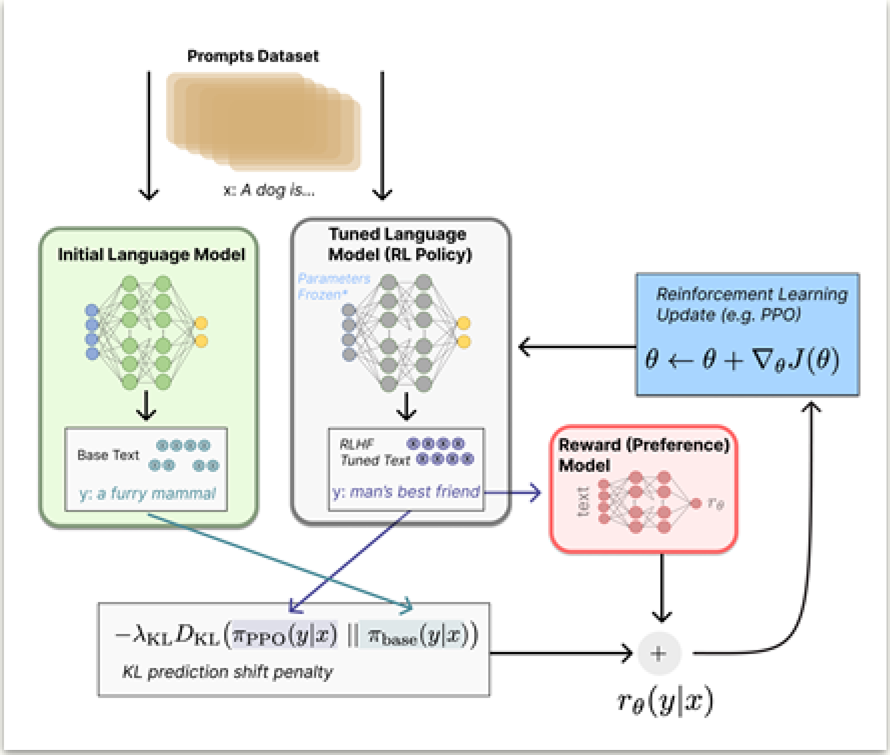
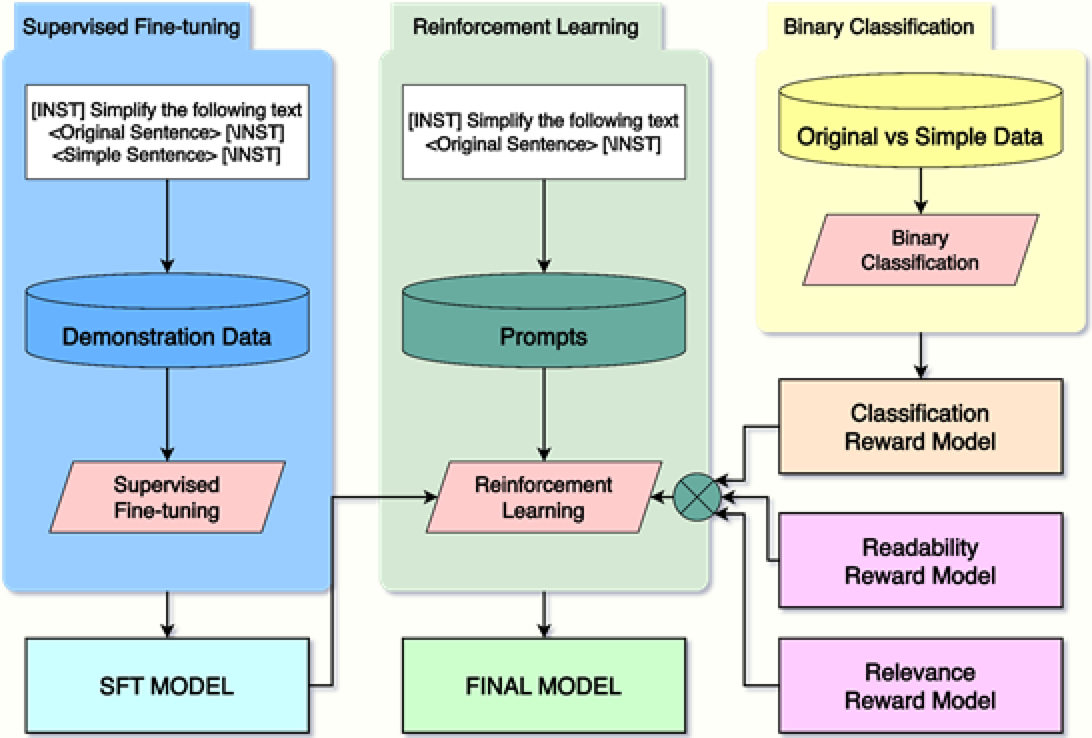


Figure from a blog by Huggingface

Automatic Text Simplification

- Rahman et al, 2024
  - Uses variant of **Llama-2**
  - Supervised Fine-tuning (SFT)
  - RL
    - Readability Reward - FKGL
    - Relevance Reward - Cosine similarity
    - Classification Reward - Predict simple or complex (RLHF component)



Rahman et al (2024), "Health Text Simplification: An Annotated Corpus for Digestive Cancer Education and Novel Strategies forReinforcement Learning"

Issues with GPT - RL solves the problem

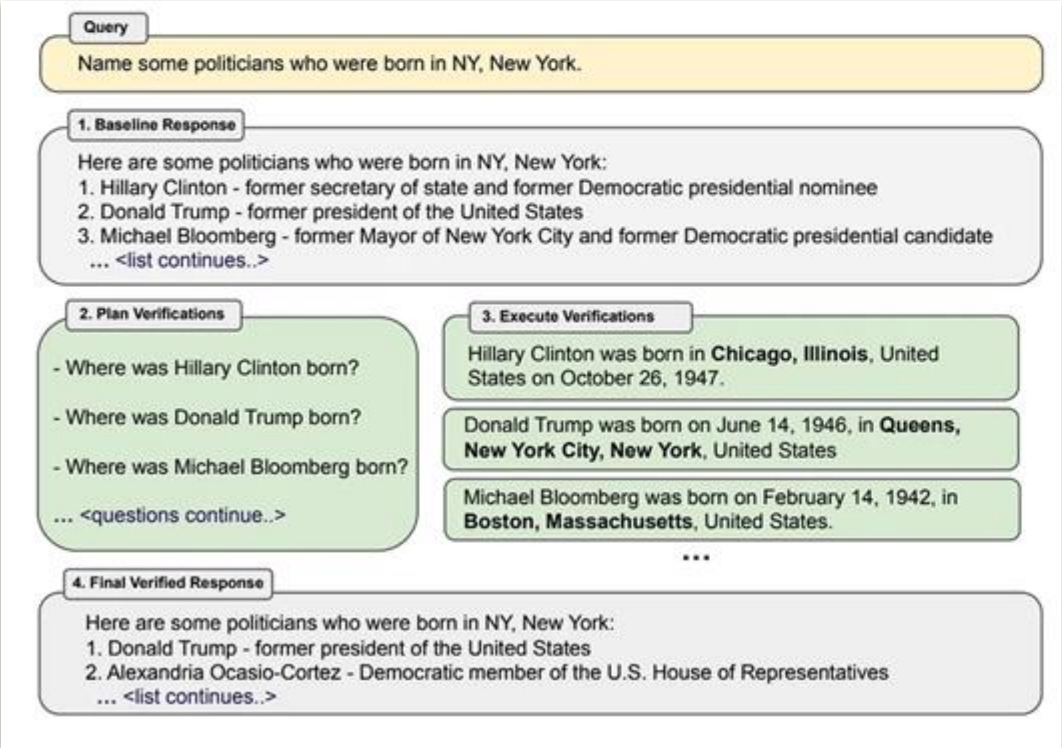
Preservation of Meaning

	Human vs Llama 2				Human vs GPT-4			
Source	Human	AI	Both	None	Human	AI	Both	None
CDC	0	4.35	95.65	0	62.32	10.14	21.74	5.80
NCI	8.89	12.22	78.89	0	56.67	15.56	27.78	0
ACS	11.11	7.19	81.70	0	50.98	6.54	41.18	1.31
Overall	8.01	8.01	83.97	0	55.13	9.94	33.02	1.93

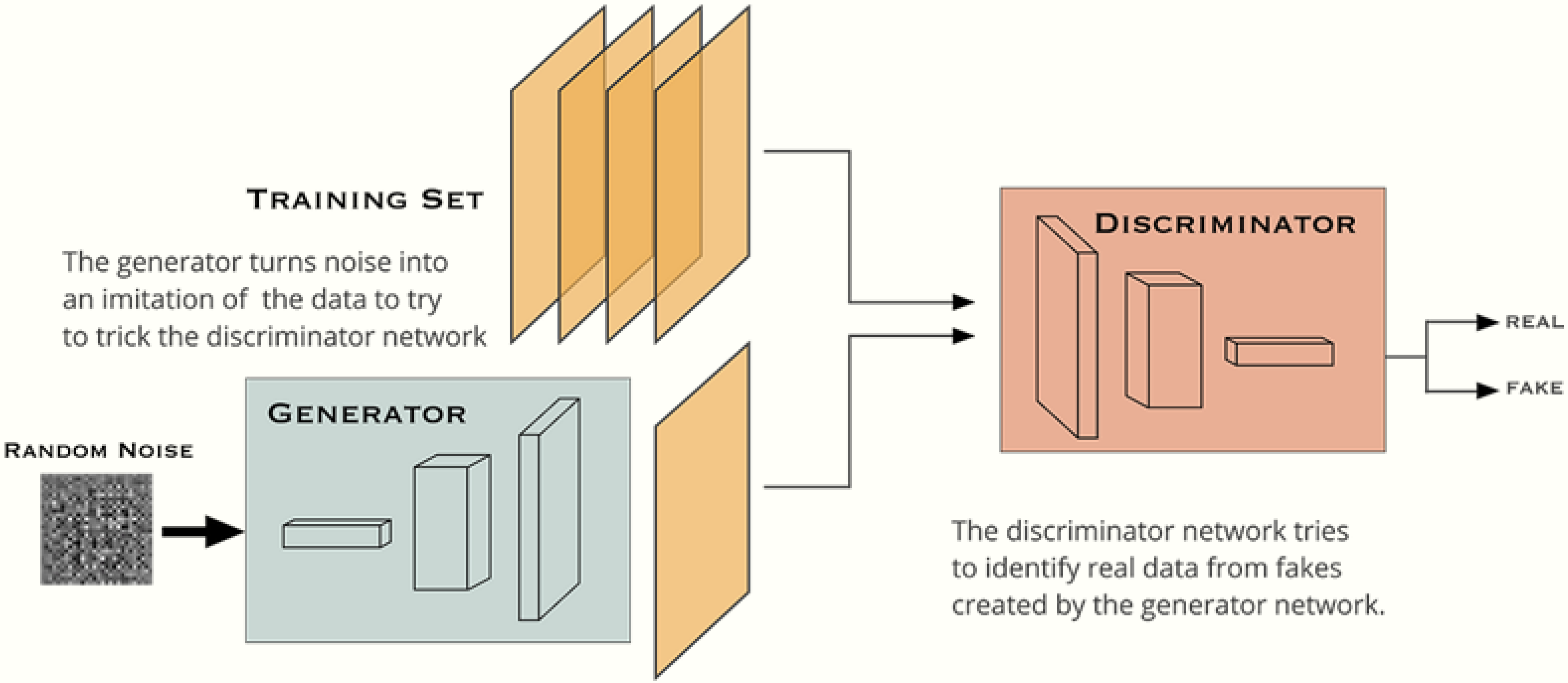
Understandability

	Human vs Llama 2				Human vs GPT-4			
Source	Human	AI	Both	None	Human	AI	Both	None
CDC	18.84	1.45	75.36	4.35	43.48	42.03	8.70	5.80
NCI	30	17.78	46.67	5.56	35.56	48.89	14.44	1.11
ACS	19.61	10.46	65.36	4.58	35.95	56.86	5.23	1.96
Overall	22.4	10.6	62.2	4.8	37.5	51.3	8.7	2.6

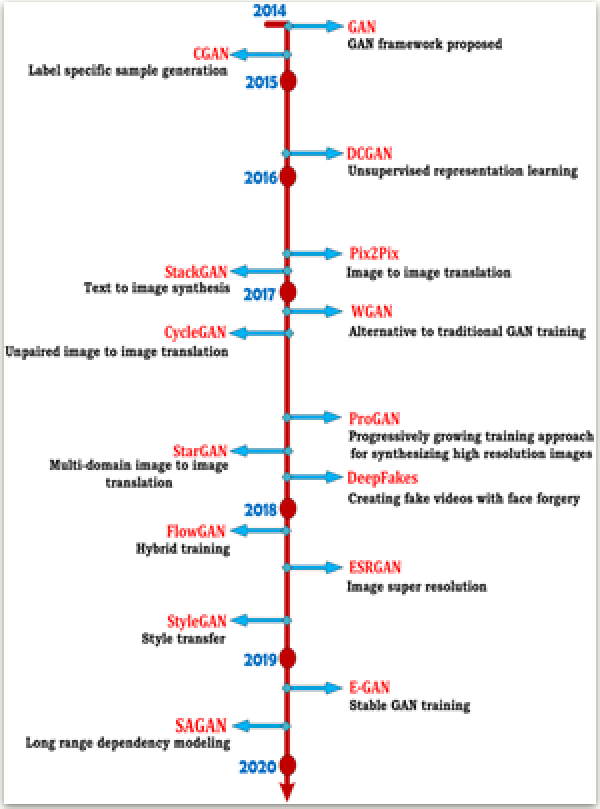
Hallucination in LLMs



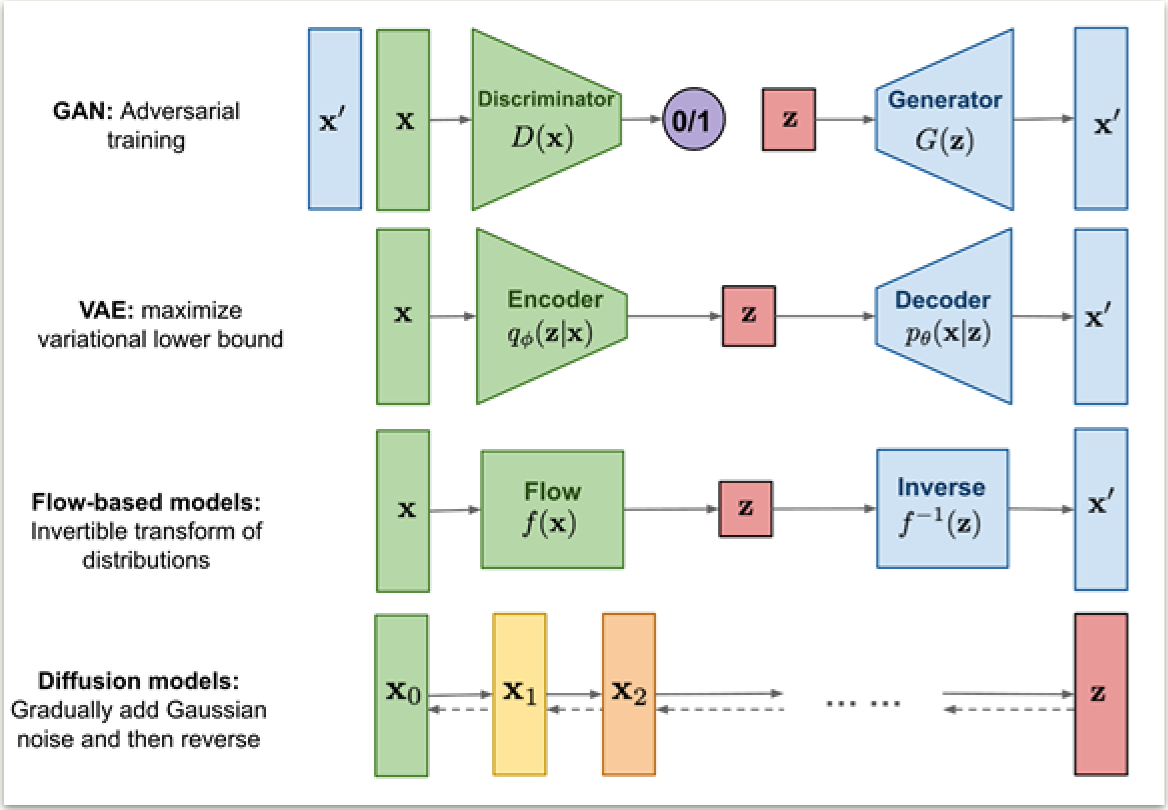
Case Study: Image Generation



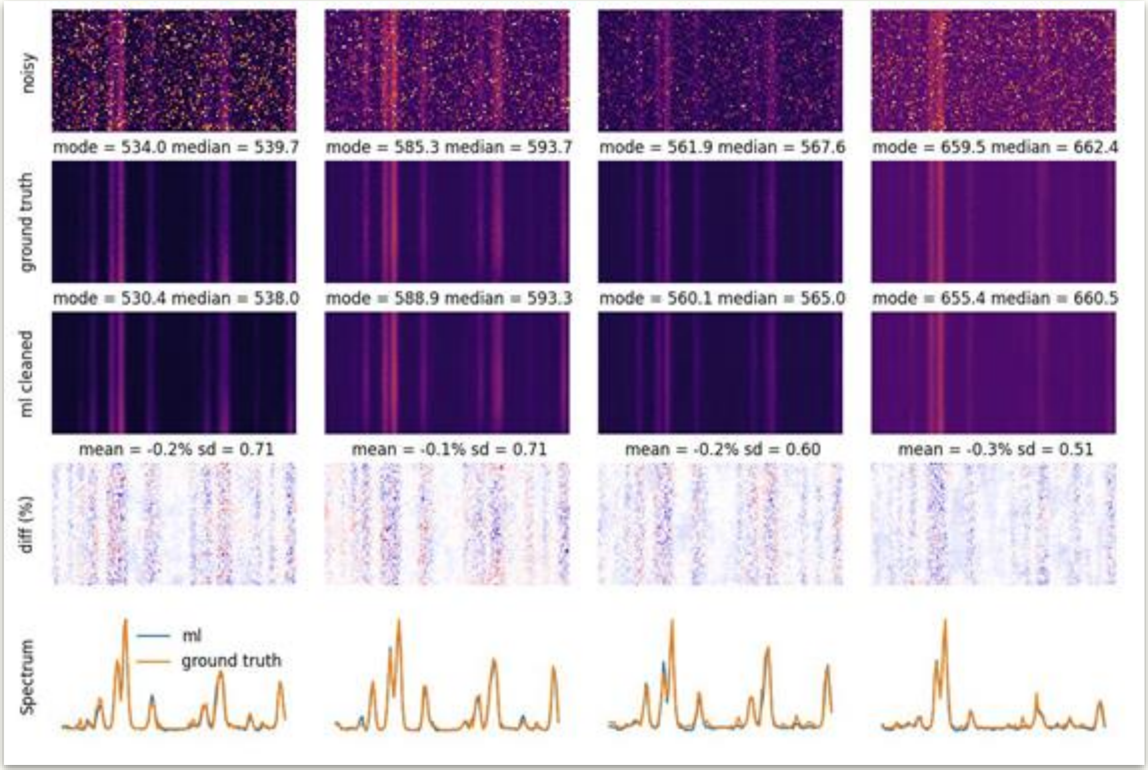
Evolution of GANs



Progress in Image Generative Algorithms



GAN in Cleaning Spectral Images



Conclusion

- ML algorithms aren’t perfect
- But can be controlled to specific need

THANK YOU