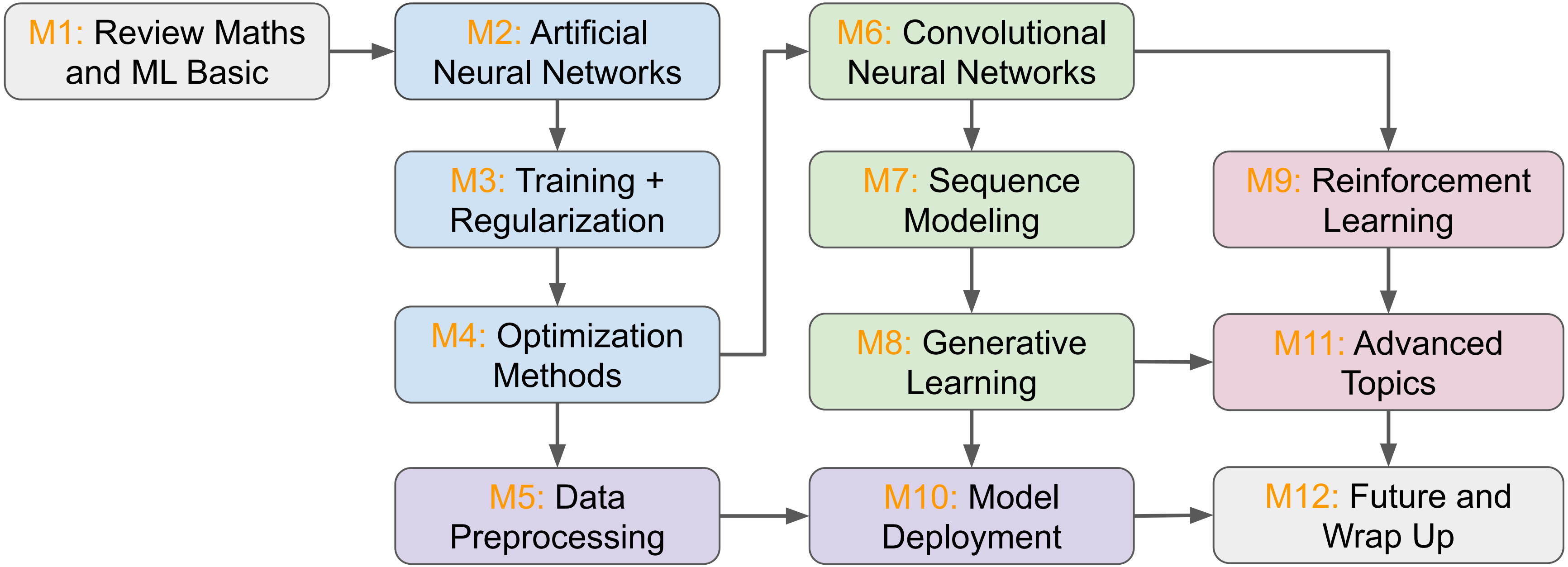
# Course Summary

## Lessons learned in SYS 6016

N. Rich Nguyen, PhD
**SYS 6016**

# SYS 6016 Roadmap

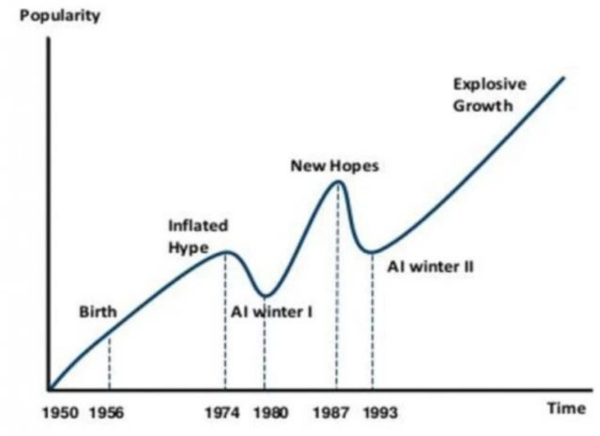


# The Rise of Deep Learning

Deep learning is an approach of AI that computes **abstract representation** in terms of other simpler representations, potentially at a multiple nested levels in a **hierarchy**.

Deep Learning has been around for a long time and known under many names. Broadly speaking, there have been three **waves of development**, separated by two **AI winters**:

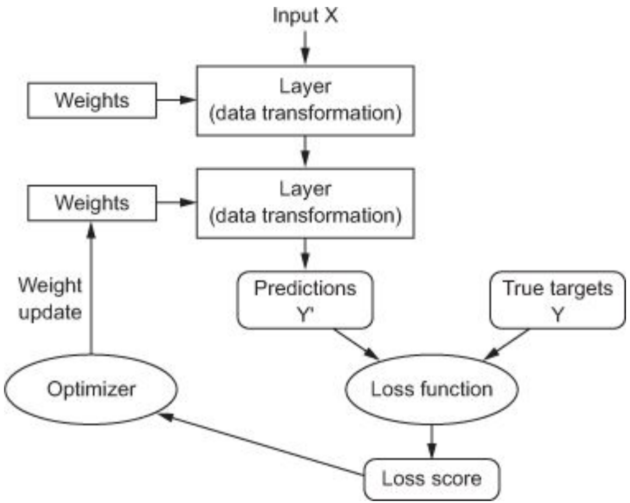
1. **Cybernetics** (1940s-1960s)
2. **Connectionism** (1980s-1990s)
3. **Deep Learning** (2006-present)



# Training a Deep Neural Network

Training a deep network revolves around the following 4 objects:

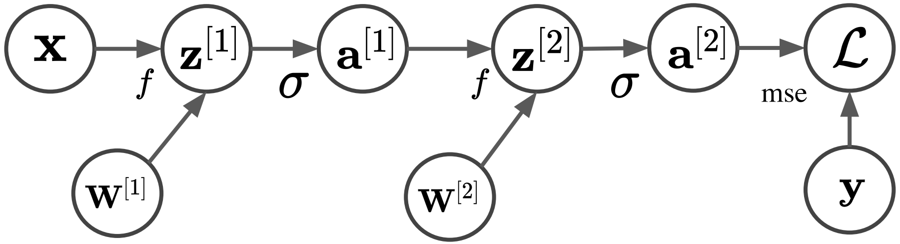
1. **Input data,** which includes features and corresponding labels
2. **Layers**, which are combined into a network
3. **Loss function**, which defines the feedback signal used for training
4. **Optimizer**, which determines how the learning proceeds



# Backpropagation Algorithm

Proposed in a 1986 groundbreaking article by **Rumelhart** and others, **back-propagation** algorithm is the most widely used algorithm in deep learning:

1. **Forward Pass:** Feed a training instance to the network, compute the output signal of every neuron, and measure the network error via a loss function
2. **Backward Pass:** From the network error at the output layer, measure (*propagate*) *the error gradient* coming from each neuron in previous layers
3. **Gradient Descent:** Run gradient descent on all connection weights using the measured error from step 2.



# Regularization Strategies

**Regularization** is the modification made to a learning algorithm that is intended to *reduce* its generalization error but not necessarily its training error. E.g., we can modify the preference for the weights to have smaller L2 norm:

$$J(\mathbf{w}) = \mathcal{L}_{\text{mse}}(\mathbf{X}_{\text{train}}, \mathbf{y}_{\text{train}}) + \lambda \mathbf{w}^\top \mathbf{w}$$

Several general strategies used to regularize neural networks:

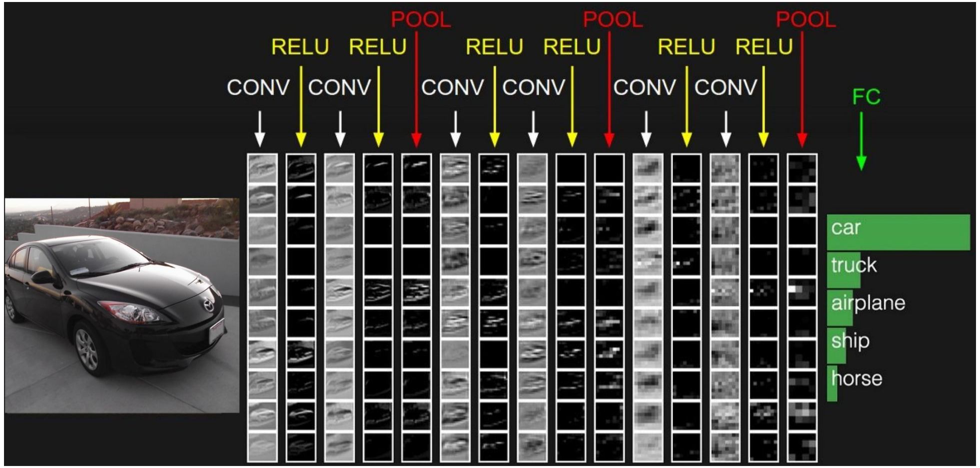
1. Parameter Norms
2. Dataset Augmentation
3. Multi-task Learning
4. Early Stopping
5. Bagging
6. Dropout
7. Tangent Prop

# Optimization Methods

1. **Initialization:** how to initialize the weights so that they do not saturate?
2. **Activation:** how to solve the vanishing gradient problem?
3. **Normalization:** how to get the model to learn the optimal scale?
4. **Optimizers:** when gradient descent was too slow or not good enough?
5. **Adaptive Learning Rate**: what if convergence is too slow or sub-optimal?
6. **Second-Order Training Methods**: can we make use of second derivatives?

# Convolutional Neural Networks

✓ Look at the **applications** of Convolutional Neural Networks (CNNs)
✓ Study the building blocks of CNNs: **convolutional** and **pooling layers**
✓ Explore winning CNN architectures: **LeNet-5**, **AlexNet**, **GoogLeNet**, **ResNet**



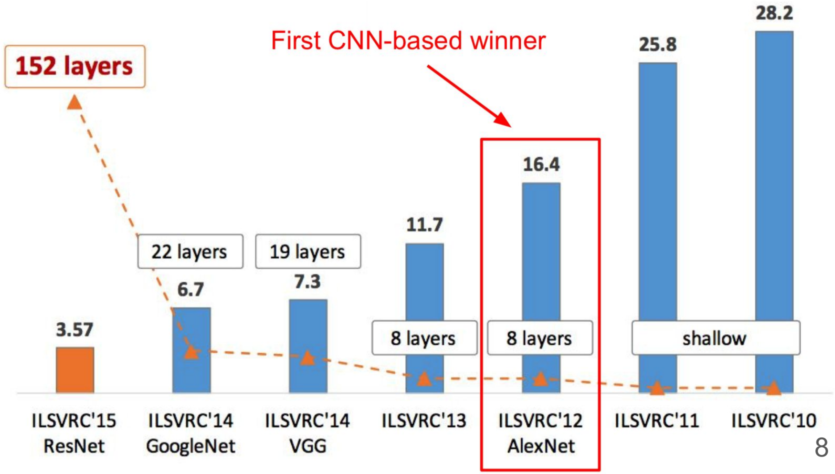
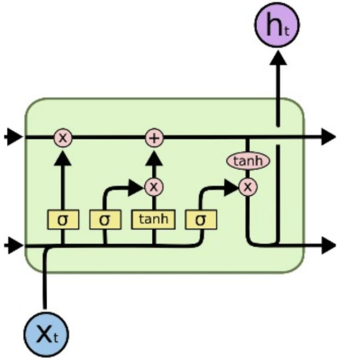






Figure copyright Kaiming He, 2016.

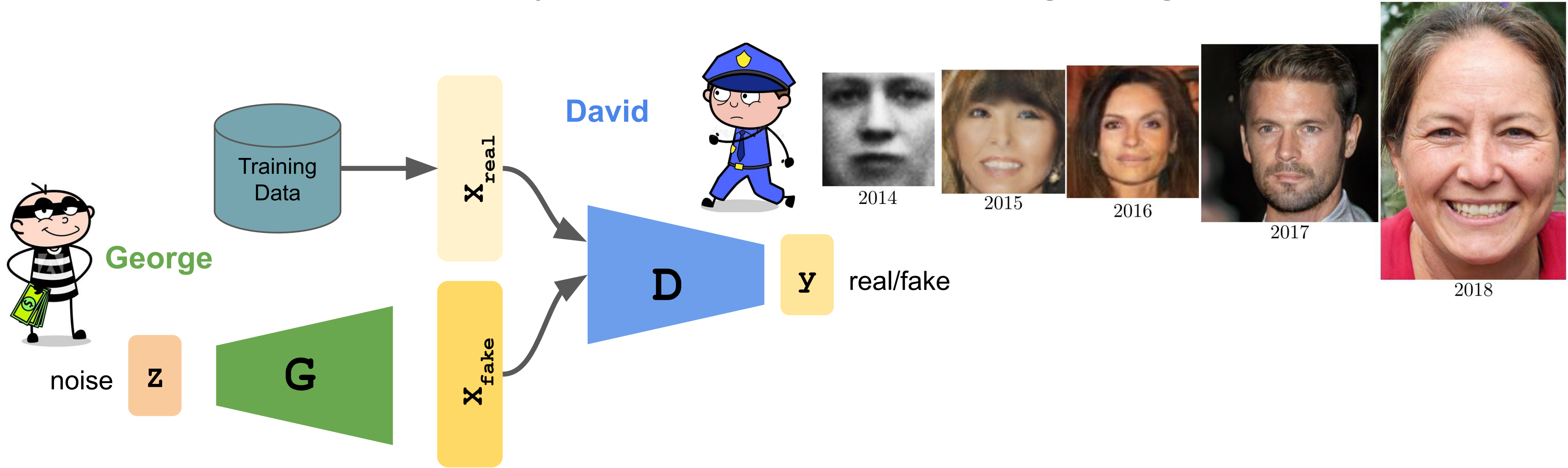
# Sequence Modeling

✓ RNNs can be used to solving sequence modeling problem
✓ LSTM maintains a separate cell state and uses gates to control the flow
✓ Attention mechanism revolutionizes machine translation and NLP
✓ Transfer Learning works! Pretrain models can be fine-tuned for new tasks
✓ Transformers and BERT are easier to train, more efficient neural networks
✓ NLP Models: Bag-of-words → RNNs / LSTM → Transformers → BERT → ?



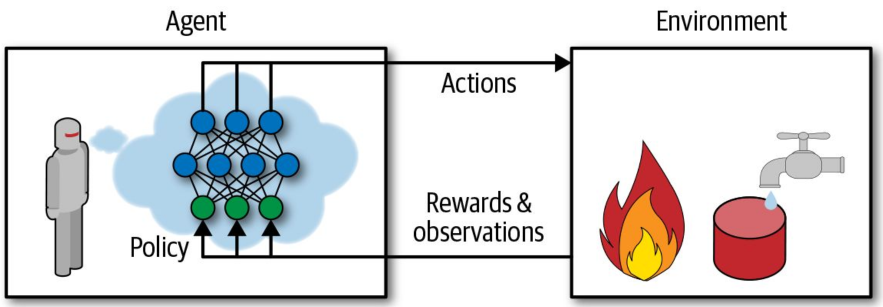

# Generative Learning

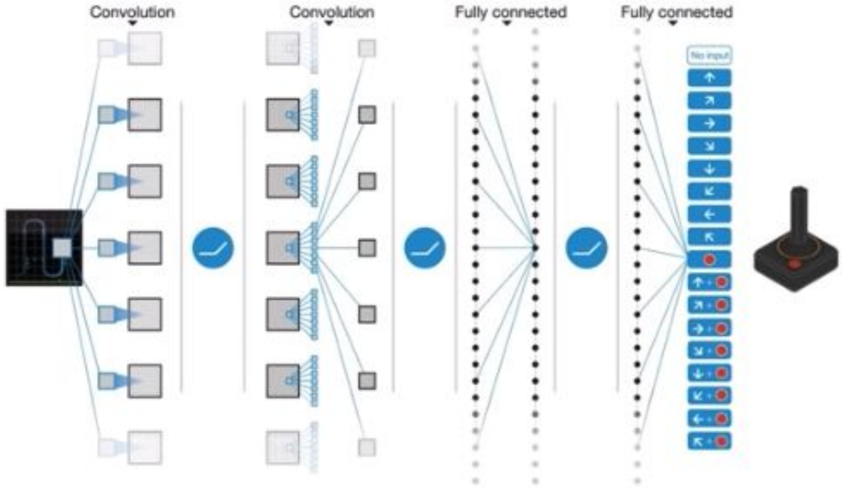
- Interpret hidden latent variables in AutoEncoder using perturbation
- GANs are “*the most interesting idea in the last ten years in ML*” -- Yann LeCun
- GANs generate realistic images that are indistinguishable from real images
- The dynamics of the adversaries are still not well understood
- There are such a variety of GANs out there and still growing



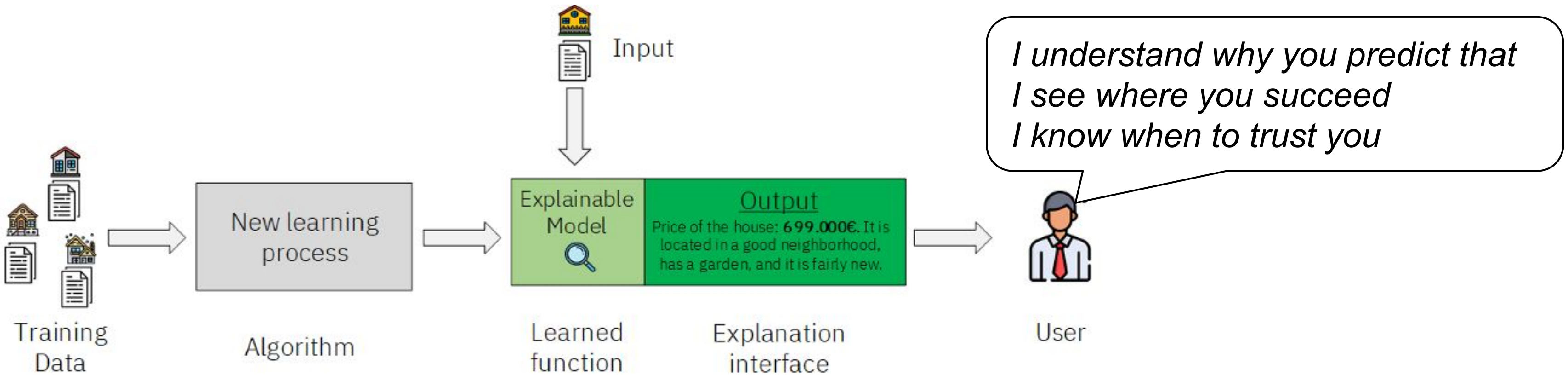
# Reinforcement Learning

- ✓ Reinforcement Learning (RL) is one of the most exciting fields of ML
- ✓ Value-based Deep Q-Networks to approximate Q function and infer action
- ✓ Policy-based Policy Gradients to learn the policy directly
- ✓ The Actor-critic algorithm and an application in AlphaGo
- ✓ Try out the TF-Agents Library



# Explainable AI (XAI)




**Goal:** Design trustworthy, reliable, and explainable AI models

There are 3 ways to design explainability into the **training process** of a model:

1. **Pre-training** explainability: understand and describe the data
2. **In-training** explainability: develop more inherently explainable models
3. **Post-training** explainability: extract explanation from trained models

Promising approaches to develop XAI **best practices**

# 4 Programming Assignments

Implementing different learning algorithms in attempt to solve a particular problem while gaining a deeper understanding of the method

ASSIGNMENT 1

**Implementing an Artificial Neural Network**

To solve the classic nonlinear XOR problem

ASSIGNMENT 2

**Recognizing UVA Historical Landmarks with a Convolutional Neural Network**

To use transfer learning of award-winning convolutional networks to classify unique buildings on Grounds

ASSIGNMENT 3

**Generating Music using Recurrent Neural Network**

To take some existing music data and train a recurrent model to learn the patterns in music that humans would be most likely to enjoy

ASSIGNMENT 4

**Applying a Deep Q-Network to play Atari Space Invaders**

To apply TF-Agents to train an agent that can achieve a superhuman level at Space Invaders

# 3 Code.a.thons

To apply what learned in the class to real-world problems

1 — 2 — 3

**Predicting Boston Housing Prices**

Find the best model given the Boston Housing dataset and improve its performance with parameter turning

**Deepfake Detection**

The spread of falsified video raises great concern. Investigate a method to deal with deepfakes called MesoNet.

**There's an app for that DL model**

Mobile phones are already a huge part of our lives. Investigate TensorFlow Lite to run TensorFlow models on mobile, embedded, and IoT devices.

# 10 Quizzes

- 10 Quizzes covering various content of the course (100 questions about various deep learning topics)
- Help you review the content each week
- Low stake (quiz each is worth only 2% of your grade)
- Allow me to gauge your understanding of the materials

# 12 Piazza Discussions

In every module, you were participating in a discussion on some advanced approaches pursued by the machine learning community.

1. Trading off bias and variance
2. Historical timeline of neural networks
3. Best design practices for deep learning models
4. Pros and Cons of optimization techniques
5. The TensorFlow dataset (TFDS)
6. Comparisons of award winning CNN frameworks
7. Pre-trained word embeddings
8. Recent advances in GANs
9. Value Learning vs. Policy Learning
10. GPU vs. Colab
11. The Downsides of Deep Learning
12. The Good, the Bad, and the Ugly of SYS 6016

# 1 Team Project: ML4VA

You were working in group of three to use deep learning to make a meaningful contribution to the well-being of the commonwealth of Virginia in eight milestones:

1. Project literature review
2. Project proposal
3. Model design and implementation
4. Project checkpoint and contribution
5. Model interpretation
6. Model deployment and scale up
7. Project presentation
8. Project report

# 1000 Grade Points (Pts)

**F**: 0-599;

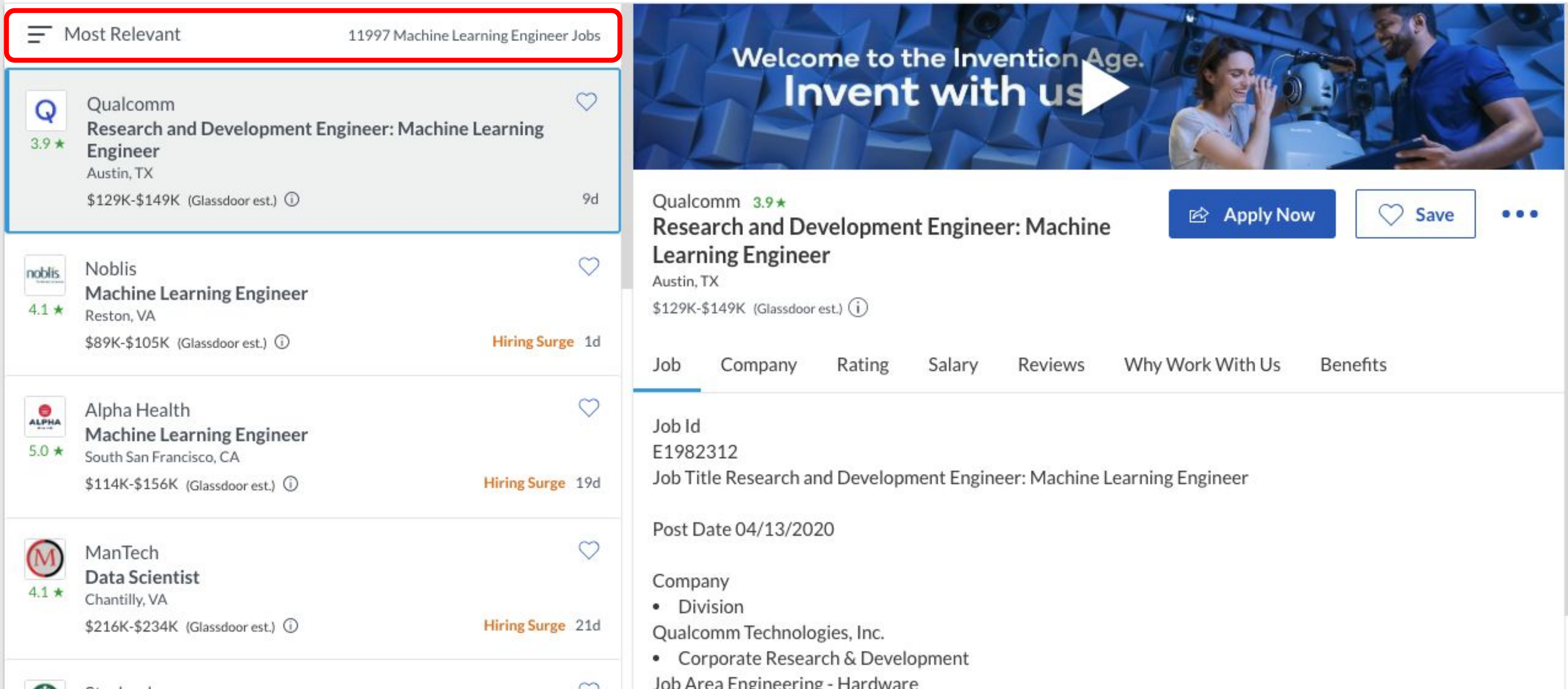
**D-**: 600-629; **D**: 630-669; **D+**: 670-699;

**C-**: 700-729; **C**: 730-769; **C+**: 770-799;

**B-**: 800-829; **B**: 830-869; **B+**: 870-899;

**A-**: 900-929; **A**: 930-969; **A+**: 970-1000

# 12,000 ML positions on Glassdoor



# Why study Machine Learning?

[your reason here]

Gain abilities to solve large-scale problems

Great opportunities for the 21st century

Obtain a data scientist position at a company

Curious to know how it works

# Thank you!

- Rich Nguyen