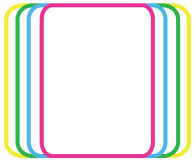
Dive into Deep Learning

A study group by dair.ai

Hello! 👋

- Hi, I am Elvis
- Educator
- Research Scientist
- Founded dair.ai
- Find me @omarsar0



dair.ai

@dair_ai

Slack group

GitHub

Agenda

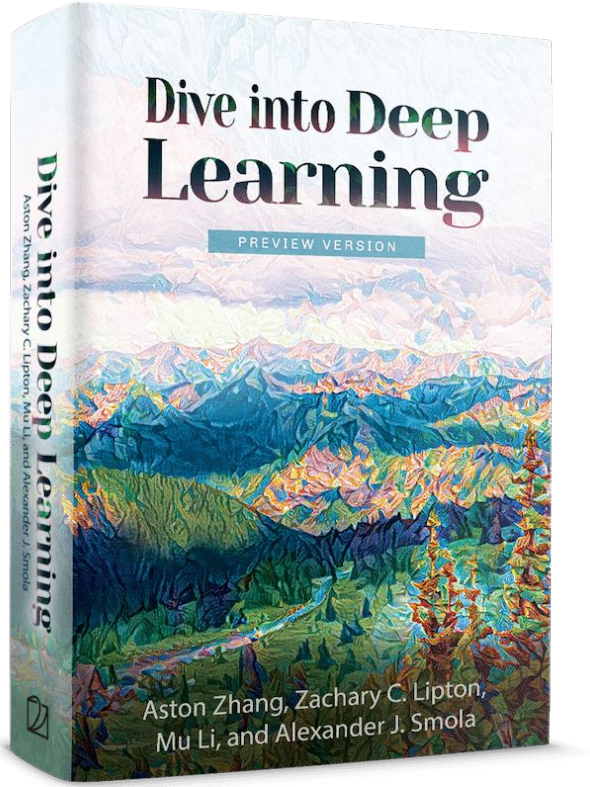
- About the study group
- Introduction to Deep Learning
- Q&A

About “Dive into Deep Learning” study group

- What it is?
  - A study group to encourage ***open discussion*** and ***participation***
- Who is it for?
  - Anyone interested to learn about ***machine learning*** and ***deep learning principles***
- What you will get out of it?
  - Learn the deep learning principles and get hands-on experience

Things to know

- We will cover the d2l.ai online book.
- The online book has over 17 chapters
    - ~17 sessions done biweekly?
- Live discussions, code walkthroughs, overviews
- Certificate of Completion
- More info in the GitHub repo



Prerequisites

- **Basics in:**
  - Linear Algebra and Calculus
  - Probability and Statistics
    - All of Statistics
  - Python Programming
    - Learn Python
    - Codecademy
  - Data Mining and Text Mining

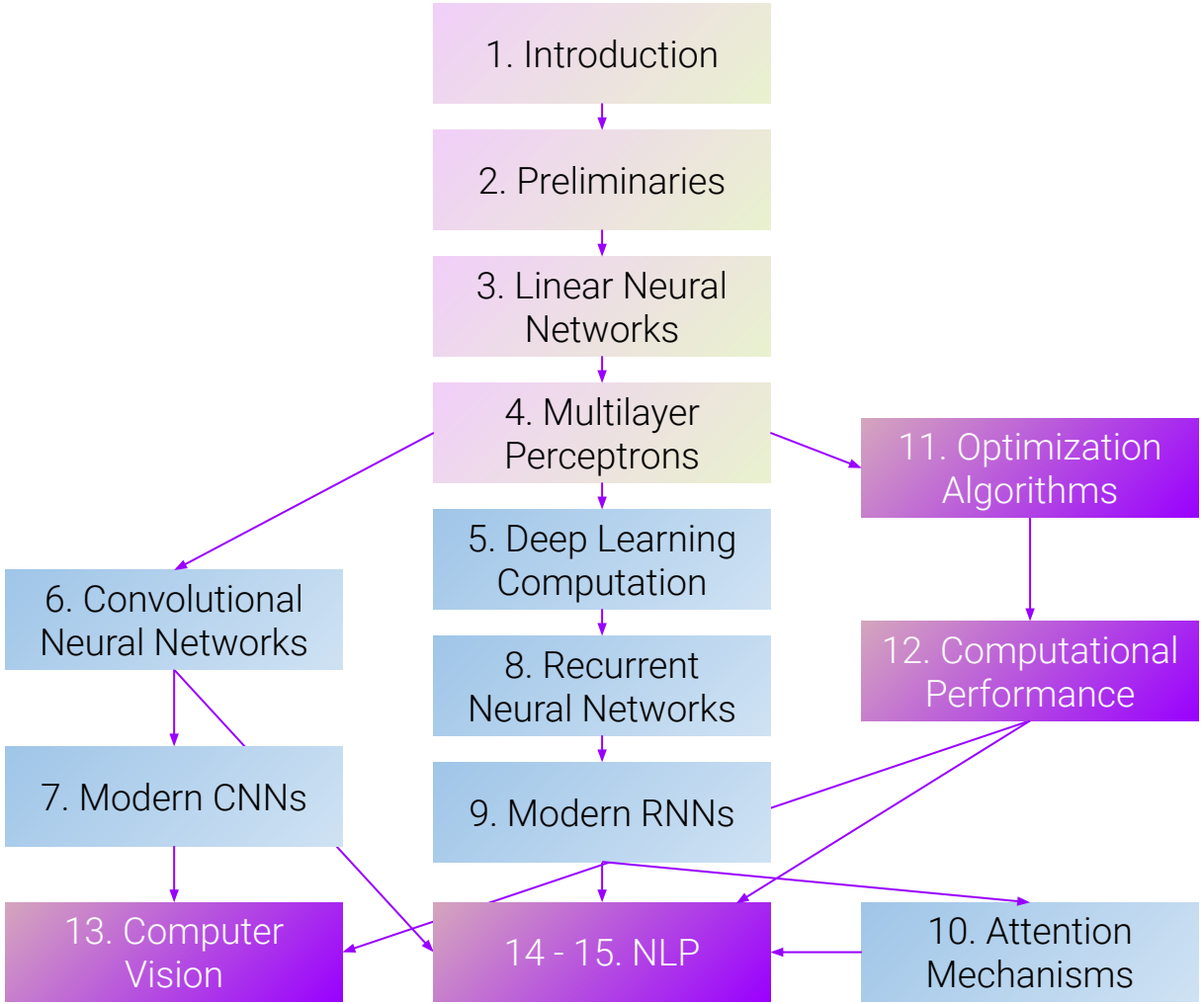
GitHub and Google Colab

- Setup a GitHub account
- Create a Google Colab Account (access to a free GPU)
- Submit code via Colab or GitHub notebook

Other important information

- Sessions are being recorded
- Use the official discussion forum for other questions
- Please take a look at assignment 1 here (due 1 week from today)
- Certificate of Completion
    - Complete at least 80% of the exercises/assignments
- Code of conduct

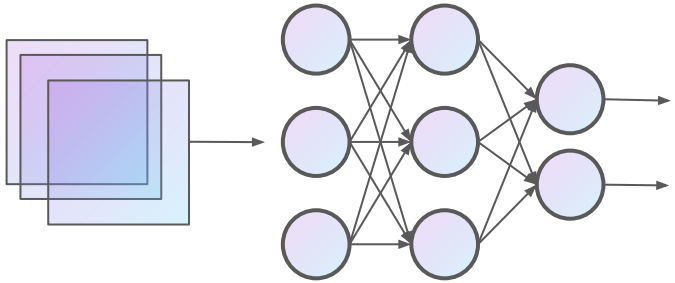
Content Structure



Introduction to Deep Learning

Deep Learning

- Neural networks were considered *outmoded tools*
- Nowadays deep learning is driving rapid progress



Reinforcement Learning

Natural Language Processing

Computer Vision

Statistical Modeling

Automatic Speech Recognition

Recent Advancements using Deep Learning

- Self-driving cars
- Draft email assistant
- Agents that play Go (beat world champion)
- Smart reply (e.g., automate customer service)
- Movie making
- Diagnosing diseases
- Physical simulators (Astrophysics)
- DNA sequencing (Biology)

dair.ai ***Introduction to Deep Learning*** (2020)

Why Machine Learning?

- You start developing with ***first principles***
- Design business logic for your application
    - Spells out which actions to take for every possible scenario
    - Before seeing a user transaction
- If you can program solutions that work 100% of the time
    - You don’t need machine learning
- When should we use machine learning?

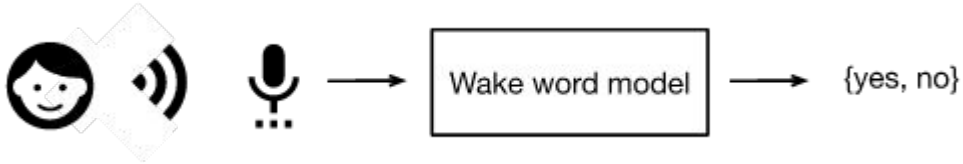
When do you need Machine Learning?

- When solving a problem requires the need to ***adapt*** the program
- ML techniques aim to ***learn from experience*** and improve performance on a ***task*** using observational data

Machine learning is the study of powerful techniques that can learn from experience

A Motivating Example

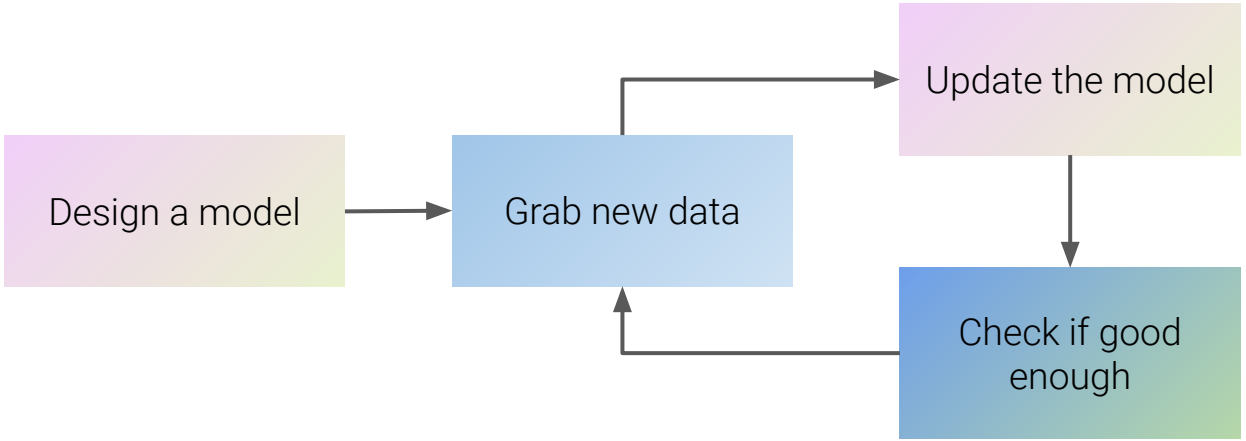
- A program that responds to the word “Alexa” or “Siri” or “Okay, Google”
- Imagine your job was to write a program to map the raw audio snippets
  - amplitude of sounds waves to predictions {yes, no}
- Can you build this from first principle?
- What makes this a hard problem?
- Would ML help?



Other Motivating Examples

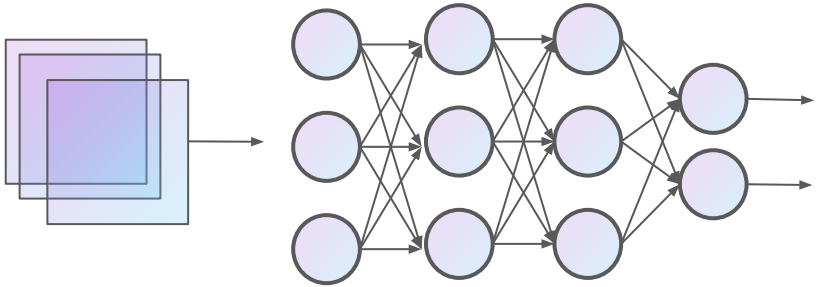
- A detector that emits large value (-/+) if cat or dog, respectively
- Real-time transcriber
- A translator of languages

Training Process



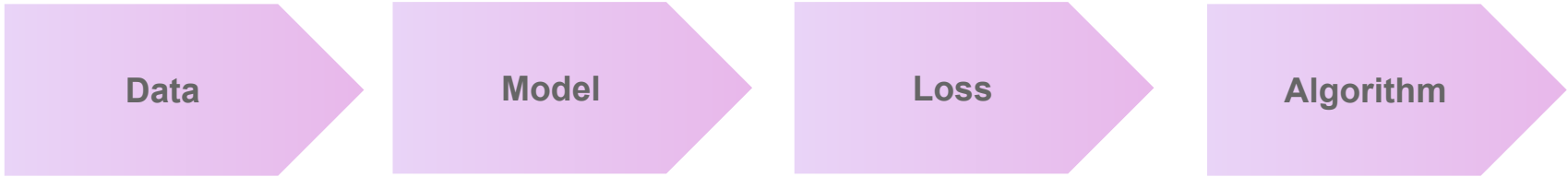
From shallow models to deep models

- In the bygone days, we ***manually-engineered features***
- With deep learning features are ***learned automatically***
- Deep learning provides a ***unified set of tools*** for tackling diverse problems that in the past required ***domain-specific preprocessing***



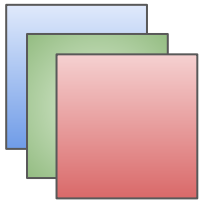
Key Components

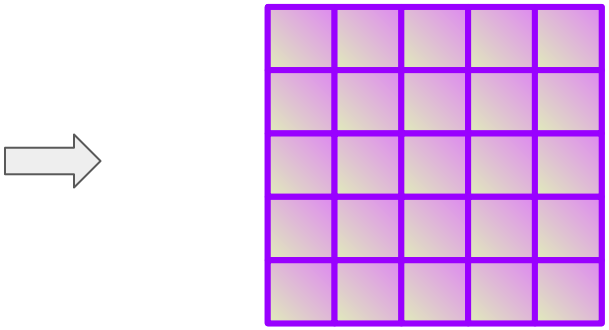
Regardless of the ML problem we have the following:

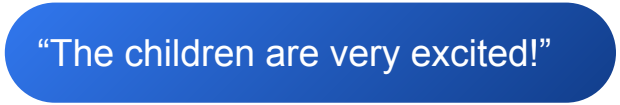


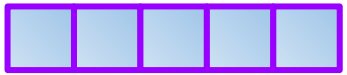
Data

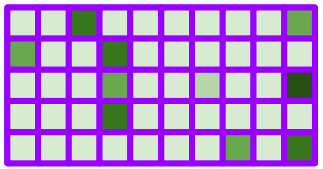
Data is a collection of examples numerically represented as features









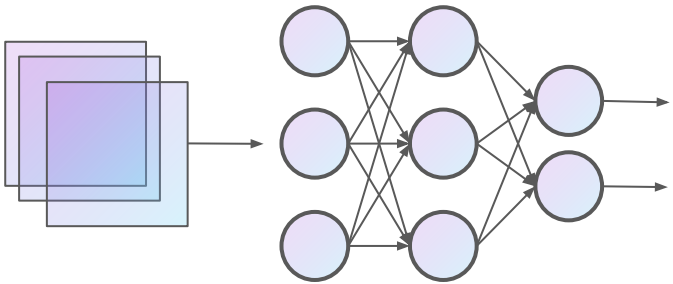


Challenges with Data

- ***Fixed length*** is a convenient property of data but it's not always present
- Data mined from the internet varies in length and needs ***standardization***
    - E.g., Images need to be cropped to standard sizes
- Text data is typically of ***varied length*** (e.g., Amazon reviews)
- Deep learning models handle ***varying-length*** data better
- With more data we can theoretically train ***more powerful models***
- Groups of people can easily be ***unrepresented*** with data
- Data can also ***reflect societal prejudices*** (e.g., automated screening resumes)

Models

- Machine learning models are ***statistical models*** that can be estimated from data
- Deep learning models ***apply many transformations*** to the data through deep layers



Objective functions

- Objective functions ***measure*** how ***good*** or ***bad*** the model is at the task
  - Typically lower is better
  - It is sometime referred to as a ***loss function*** or ***cost function***
- Loss function defined w.r.t to the ***model's parameters***
  - The aim is to find the best parameters to ***minimize the loss*** given training data
- Training error vs. test error
  - Check for generalization and overfitting

Optimization algorithms

- The algorithm used to search for the ***best possible parameters*** that minimize the loss function
- In deep learning these algorithms are typically based on ***gradient descent***
    - Batch gradient descent, SGD, mini-batch

Momentum

Adagrad

Adam

AMSGrad

Adam

Sebastian Ruder ***An overview of gradient descent optimization algorithms*** (2016)

Kinds of ML

- **Supervised ML**
  - Regression | Classification | Tagging | Search and ranking | Recommender systems | Sequence learning
- **Unsupervised ML**
  - Clustering | PCA | Representation learning | Probabilistic graphical models | GANs
- **Interacting with Environment**
  - From predictive models to intelligent agent taking actions in environment
- **Reinforcement learning** (generalization of previous)

Roots

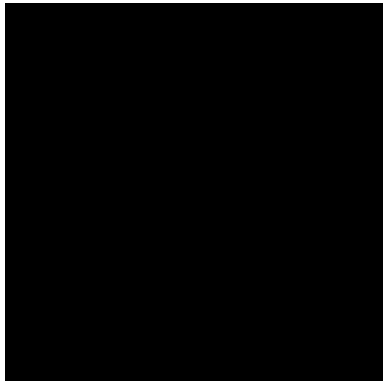
- ***Predictive capabilities*** has been desired for centuries
  - Bernoulli/Gaussian distribution used in ***natural sciences***
- ***Estimation*** used in the middle ages for obtaining average foot length
- With more ***data availability***, statistics took off
  - Linear Discriminant Analysis (dimensionality reduction)
- ***Theory of Computation*** and ***Information Theory*** both influenced ML
- ***Neuroscience*** and ***Psychology*** also influenced ML development

Deep learning *involves alternating linear and non-linear transformations combined with backpropagation for adjusting parameters on the entire network at once*

Road to Deep Learning

- ***Big data*** and ***cheaper computation***
- ***Memory efficiency*** via nonlinearities (squashing)
- Some key ideas:
    - Dropout | Attention mechanism | Memory networks | GANs | distributed training | Simulation in RL through parallelized computation
- Let's not forget the importance of ***deep learning frameworks***
    - Caffe | Torch | Theano | TensorFlow | Apache MxNet | PyTorch | Jax | ...

Success stories



Agent playing Breakout

- Intelligent digital assistants
- Object recognition
- Intelligent game-playing agents
- Self-driving cars
- Robotics, logistics, computational biology, particle physics, astronomy,...
- ML is becoming a ubiquitous tool for engineers and scientists
- Usability of AI systems in the real world becomes important
- Beginning to discuss/address fairness, racial, gender, age bias

What’s next?

- **Session 2: Preliminaries**
  - Probability, random variables, Bayes’ theorem, expectation and variance
  - I will provide extra readings and additional exercises to practise
  - Presentation + Code walkthrough
- I will announce next date/time on our GitHub repo and our Slack group

Q&A