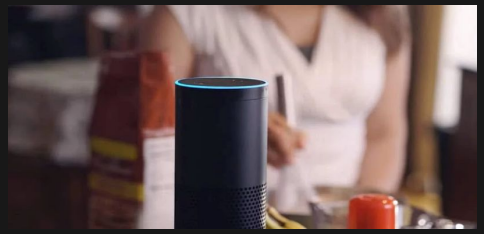
Machine Learning
Deep Learning

siddrrsh@gmail.com
ronithganjigunta@gmail.com

Self driving cars



Intelligent Agents

Google Translate

Machine Translation

Face Recognition

Medical diagnosis

Predict SAT scores?

Predict who will win the NBA championship this year?

Basics of Machine Learning

Machine Learning = design programs with the ability to improve as they acquire more experience

Data

Algorithm

Model

Machine Learning

ML algorithms discover patterns in data, and construct models using these discoveries.

These models are then used to make predictions on unseen data.

ML Terminology

- Labels - the thing we're predicting
  - the y variable in simple linear regression.
- Features - an input variable
  - the x variable in simple linear regression

An **example** is a particular instance of data, **x**.

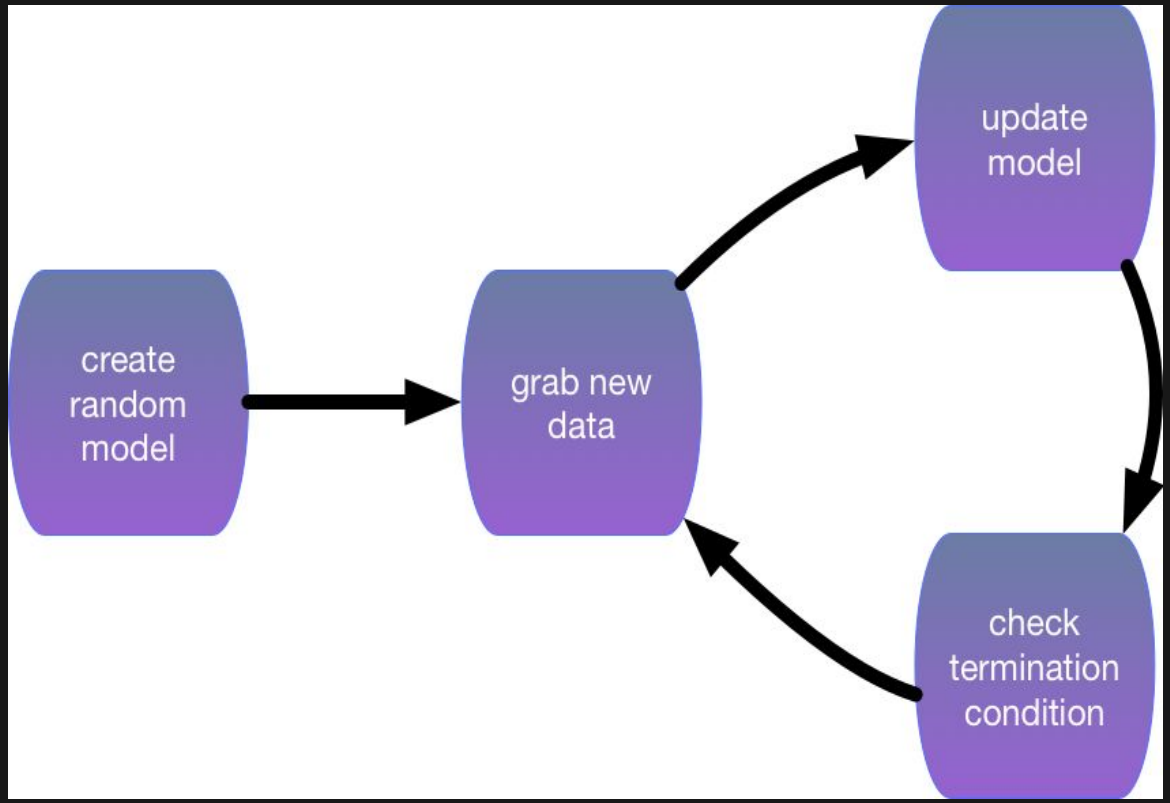
- labeled examples
- unlabeled examples

ML Terminology (Continued)

Models - Defines the relationship between features and label.

- **Training** means creating or **learning** the model (show the model labeled examples and enable the model to gradually learn)
- **Inference** means applying the trained model to unlabeled examples

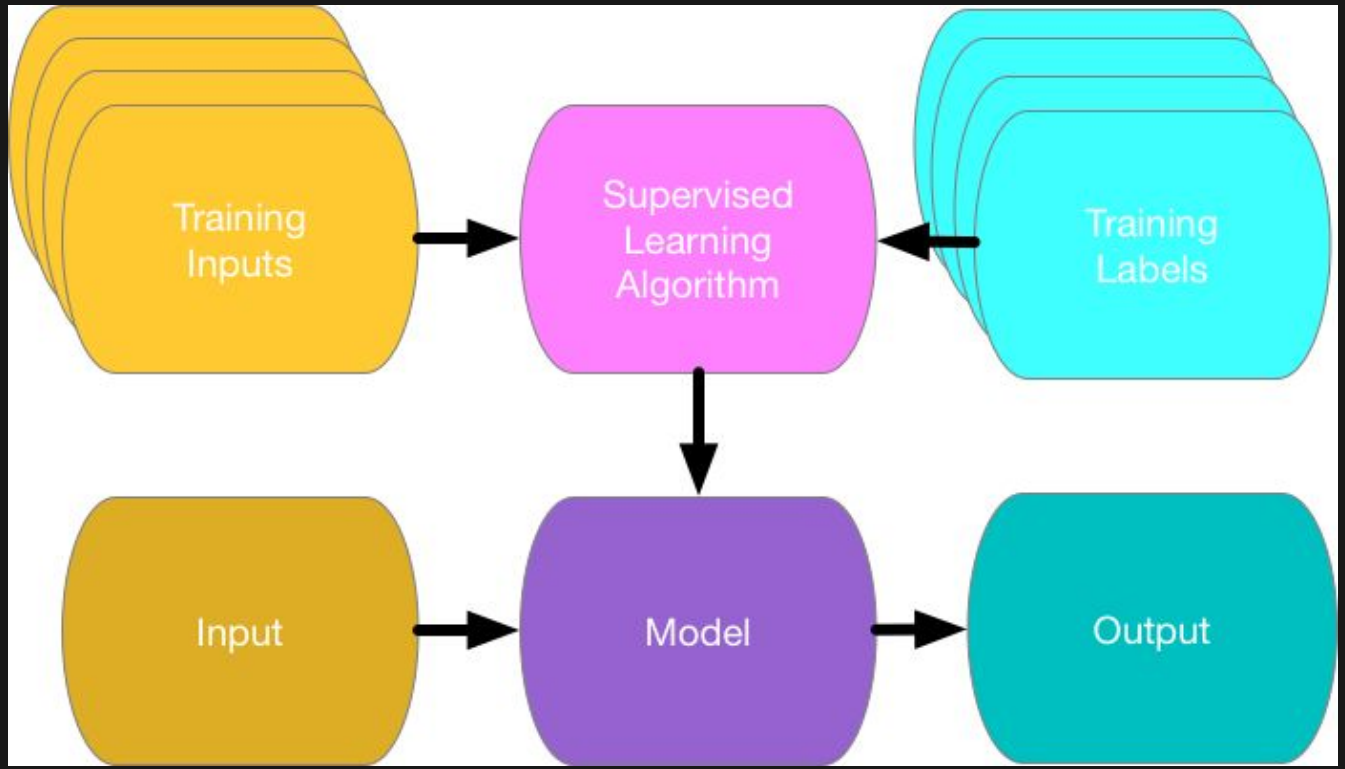
Creating a Model



Supervised Learning

- Supervised learning addresses the task of predicting targets given input data.
- The goal is to produce a model $f\theta$ that maps an input $\mathbf{x}$ to a prediction $f\theta(\mathbf{x})$
  - **Predict cancer vs not cancer, given a CT image.**
  - **Predict the correct translation in French, given a sentence in English.**
  - **Predict the price of a stock next month based on this month's financial reporting data.**

Supervised Learning (Cont'd)



Loss Functions

- Loss functions give us a way of measuring how *“bad”* our output/prediction is
- Typically the *learning* part of machine learning consists of minimizing this loss function.
  - training vs. test data
- We will cover more on this in future lectures

Optimization Algorithms

- Need some way of taking the model and its loss functions, and searching for a set of parameters that minimizes the loss
- Gradient descent and stochastic gradient descent
- In short, they look to see, for each parameter which way the training set loss would move if you jiggled the parameter a little bit.

Regression

Predicts continuous values

What is the value of a house in California?

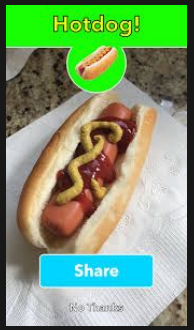
How many cars are there in this photo?

Classification

Predicts discrete values

Hot dog or not hot dog?

Is this movie a romantic comedy, documentary, or thriller?



Try this: https://itunes.apple.com/us/app/not-hotdog/id1212457521?mt=8

Pop Quiz: Regression or Classification

Is this an image of a dog, a cat, or a hamster?

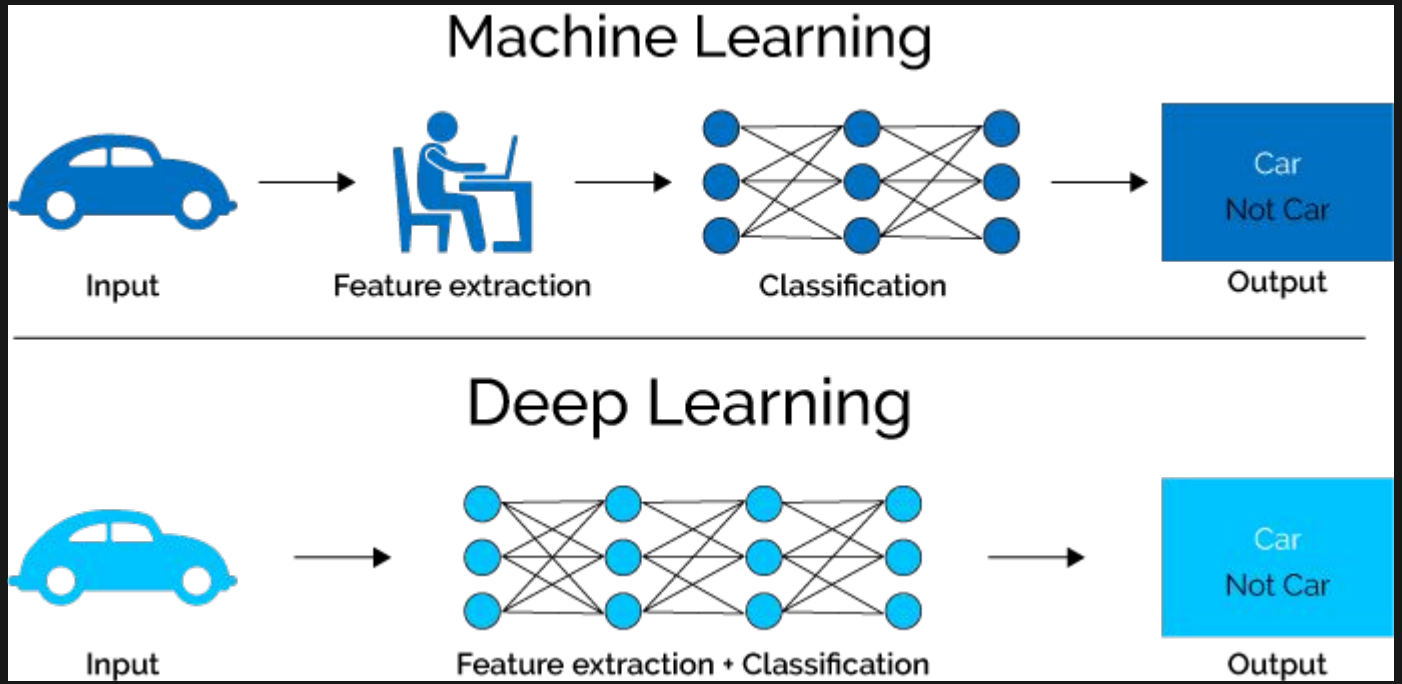
What is the probability that a user will click on this ad?

Is a given email message spam or not spam?

What is the probability a customer will buy this product or not buy this product?

Which category of products is most interesting to this customer?

Deep Learning



Deep Learning Basics

- Imitate how the human brain can process light and sound stimuli into vision and hearing.
  - Consists of multiple layers made up of hardware and GPUs.
- Cascade of nonlinear processing unit layers in order to extract or transform features of the data.
  - The output of one layer serves as the input of the successive layer.
- Absorbs the most data and has been able to beat humans in some cognitive tasks.
  - Computer vision and speech recognition have both realized significant advances from deep learning approaches.

Tagging and Parsing (TTS)

- Treat as a binary classification problem (“either both a cat and dog, or neither a cat and dog”)
- Related to computer vision

Practice Problems

Suppose you want to develop a supervised machine learning model to predict whether a given email is "spam" or "not spam." Which of the following statements are true?

- The labels applied to some examples might be unreliable.
- Emails not marked as "spam" or "not spam" are unlabeled examples.
- Words in the subject header will make good labels.
- We'll use unlabeled examples to train the model.

Suppose you want to develop a supervised machine learning model to predict whether a given email is "spam" or "not spam." Which of the following statements are true?

We'll use unlabeled examples to train the model.

✓ Emails not marked as "spam" or "not spam" are unlabeled examples.

Because our label consists of the values "spam" and "not spam", any email not yet marked as spam or not spam is an unlabeled example.

2 of 2 correct answers.

✓ The labels applied to some examples might be unreliable.

Definitely. It's important to check how reliable your data is. The labels for this dataset probably come from email users who mark particular email messages as spam. Since most users do not mark every suspicious email message as spam, we may have trouble knowing whether an email is spam. Furthermore, spammers could intentionally poison our model by providing faulty labels.

1 of 2 correct answers.

Words in the subject header will make good labels.

Suppose an online shoe store wants to create a supervised ML model that will provide personalized shoe recommendations to users. That is, the model will recommend certain pairs of shoes to Marty and different pairs of shoes to Janet. Which of the following statements are true?

- Shoe size is a useful feature.
- User clicks on a shoe's description is a useful label.
- The shoes that a user adores is a useful label.
- Shoe beauty is a useful feature.

Suppose an online shoe store wants to create a supervised ML model that will provide personalized shoe recommendations to users. That is, the model will recommend certain pairs of shoes to Marty and different pairs of shoes to Janet. Which of the following statements are true?

✓ Shoe size is a useful feature.

Shoe size is a quantifiable signal that likely has a strong impact on whether the user will like the recommended shoes. For example, if Marty wears size 9, the model shouldn't recommend size 7 shoes.

1 of 2 correct answers.

✓ User clicks on a shoe's description is a useful label.

Users probably only want to read more about those shoes that they like. User clicks is, therefore, an observable, quantifiable metric that could serve as a good training label.

2 of 2 correct answers.

The shoes that a user adores is a useful label.

Shoe beauty is a useful feature.

What you need to know?

Linear Algebra: Matrices and Vectors
(We will cover this in the future)

Probability and Statistics

(Linear Regression)

Programming: Python

Logic and computational skills

Proposed schedule (subject to change):

09/12: Lecture: Introduction to Machine Learning and AI
09/26: Lecture: Regressions/Classifications and Loss Functions
10/10: Hands-on: Regressions and Classifications
10/24: Guest Speaker
11/14: Lecture: Deep Learning and Introduction to Neural Nets
11/28: No class (Thanksgiving)
12/12: Lecture: Perceptron and deep neural networks
01/09: Lecture: Deep Learning and Computer Vision/NLP
01/23: Hands-on: Introduction to TensorFlow and MXnet
02/13: Guest Speaker
02/27: Hands-on: Introduction to Keras and Gluon
03/13: Hands-on: Train a convolutional neural network
03/27: **last meeting**

Software to download:

- Python, Anaconda Version 4
- Sublime Text or a text editor
- Github
- Tensorflow Libraries (numpy, scipy)

Next Meeting:

- Linear Regression and Classifications
- Working with Loss Functions

Resources

Stanford ML
https://www.coursera.org/learn/machine-learning/home/welcome

Google ML crash course
https://developers.google.com/machine-learning/crash-course/ml-intro

Princeton Deep Learning
https://www.cs.princeton.edu/courses/archive/spring16/cos495/slides/Intro_to_course.pdf

Glu-on tutorial
https://gluon.mxnet.io/chapter01_crashcourse/introduction.html#Basics-of-machine-learning

Videos (if time permits)

AI in a nutshell
https://www.youtube.com/watch?v=mJeNghZXtMo

Google Machine Learning
https://www.youtube.com/watch?v=nKW8Ndu7Mjw

AI applications
https://www.youtube.com/watch?v=GapiDqifthM

Simplified neural nets
https://www.youtube.com/watch?v=rEDzUT3ymw4