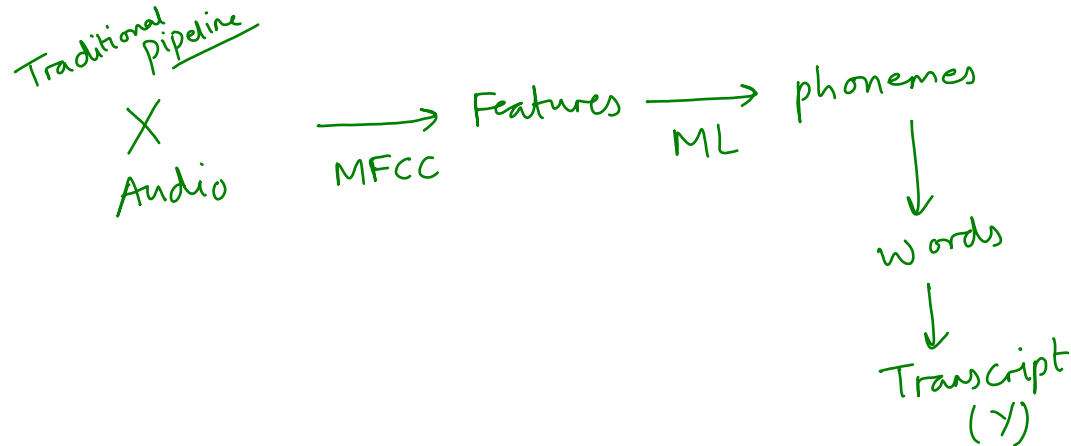End-to-end deep learning

---

What is
end-to-end
deep learning

What is end-to-end learning?

Speech recognition example

Traditional Pipeline

X
Audio $\xrightarrow[\text{MFCC}]{}$ Features $\xrightarrow[\text{ML}]{}$ phonemes $\downarrow$ Words $\downarrow$ Transcript (Y)

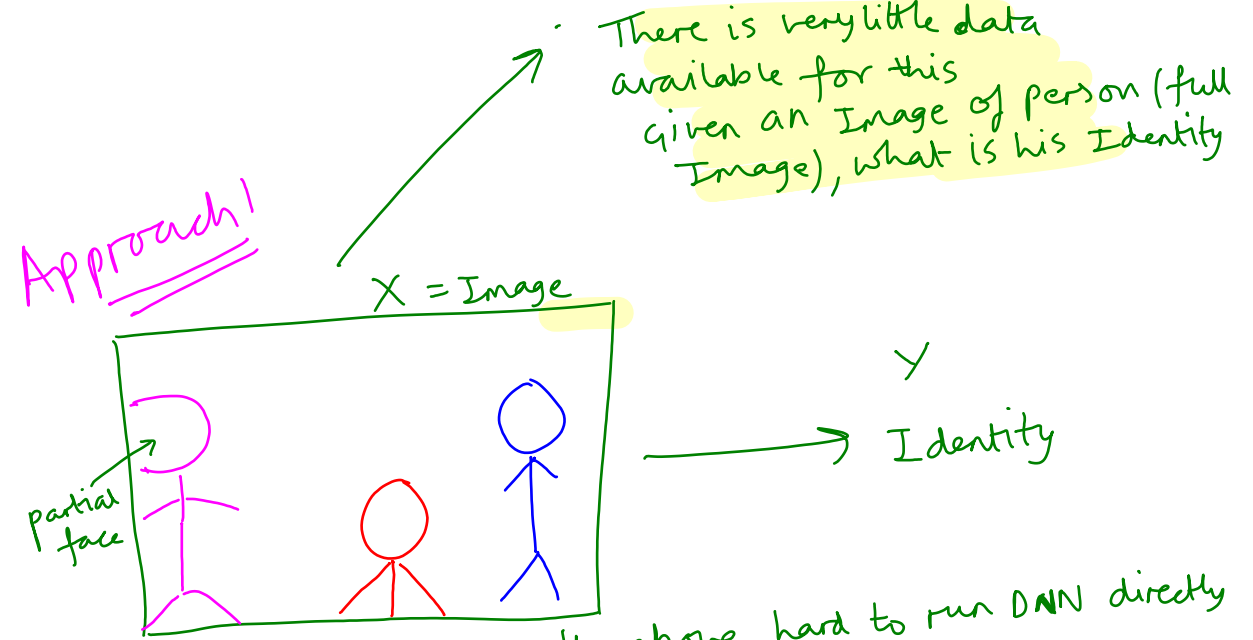— E2E Deep learning

— Audio $\xrightarrow{}$ Transcript
(X)   (Y)

— only works on very large
datasets (10K Hours to 100K Hours)
won't work on "small" datasets,
          say 3000 hours
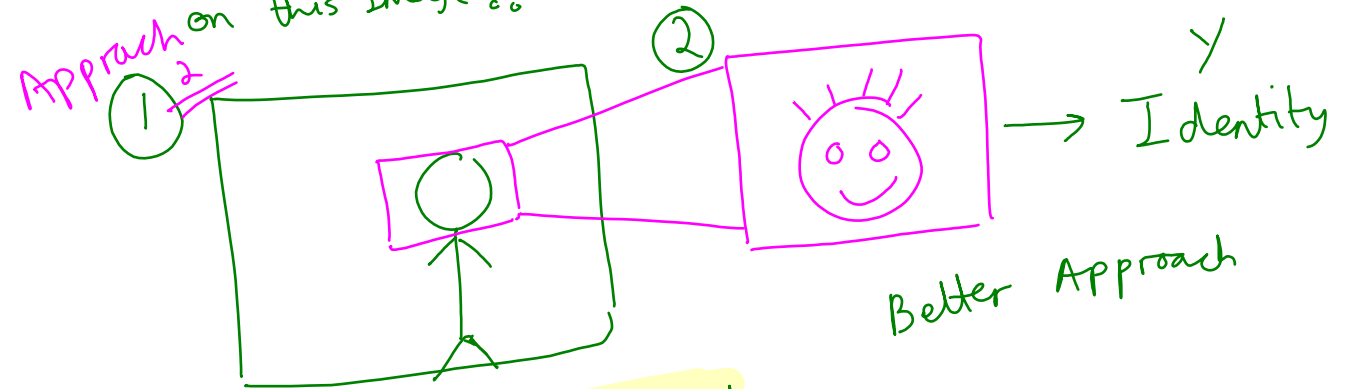— For small dataset, use traditional pipeline

- There may be multiple NN
  that you need to train to
  solve diff stages of a Deep
  learning Problem

- Deep E2E learning helps give a single
  DNN for that

Andrew Ng

Face recognition

[Image courtesy of Baidu]

Face Recognition → instead of RFID for entry

There is very little data available for this. Given an Image of person (full Image), what is his Identity

Approach 1

X = Image

partial face

Y → Identity

The i/p could be Any of the above, hard to run DNN directly on this Image ∴ extract face, then apply DNN

Approach 2

① ②

Y → Identity

Better Approach

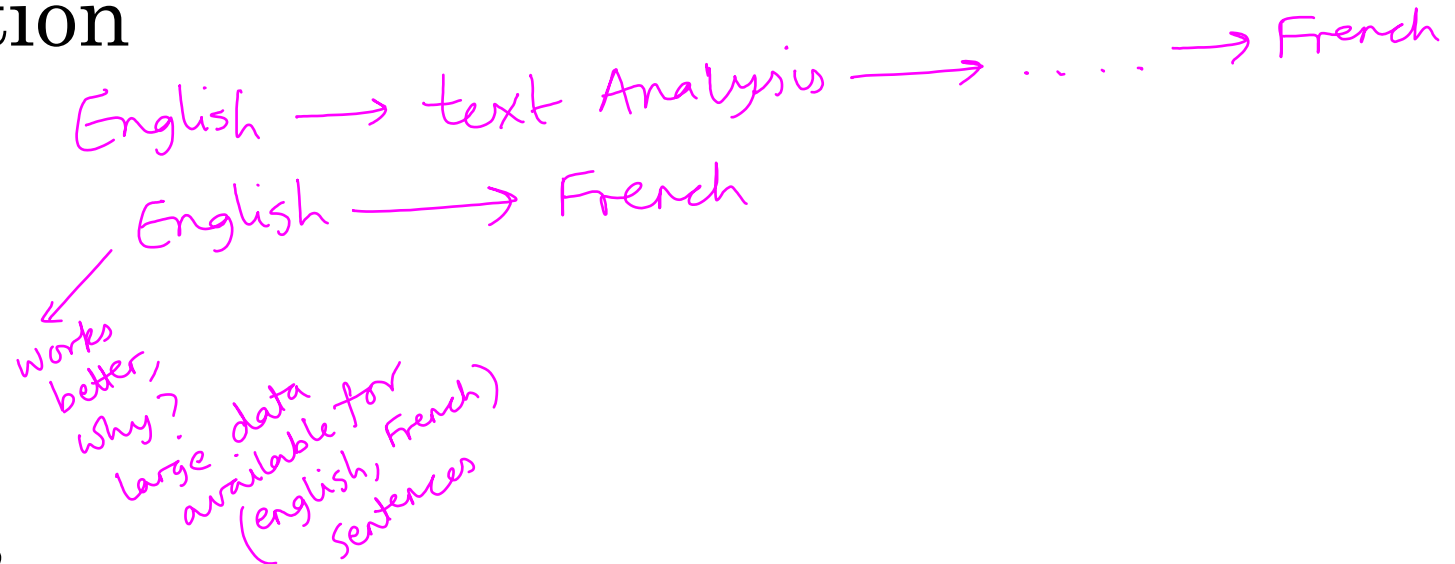— why does 2 step process work better?
You have a lot of data for each task

① Given Image Find location of face (DNN)

② Given face, match his Image to employee DB (DNN)

Andrew Ng

More examples

Machine translation

English $\longrightarrow$ text Analysis $\longrightarrow$ . . . . $\longrightarrow$ French

English $\longrightarrow$ French

works better, why?

large data available for (english, French) sentences

Estimating child's age: