HPC- Tech skills learning final output

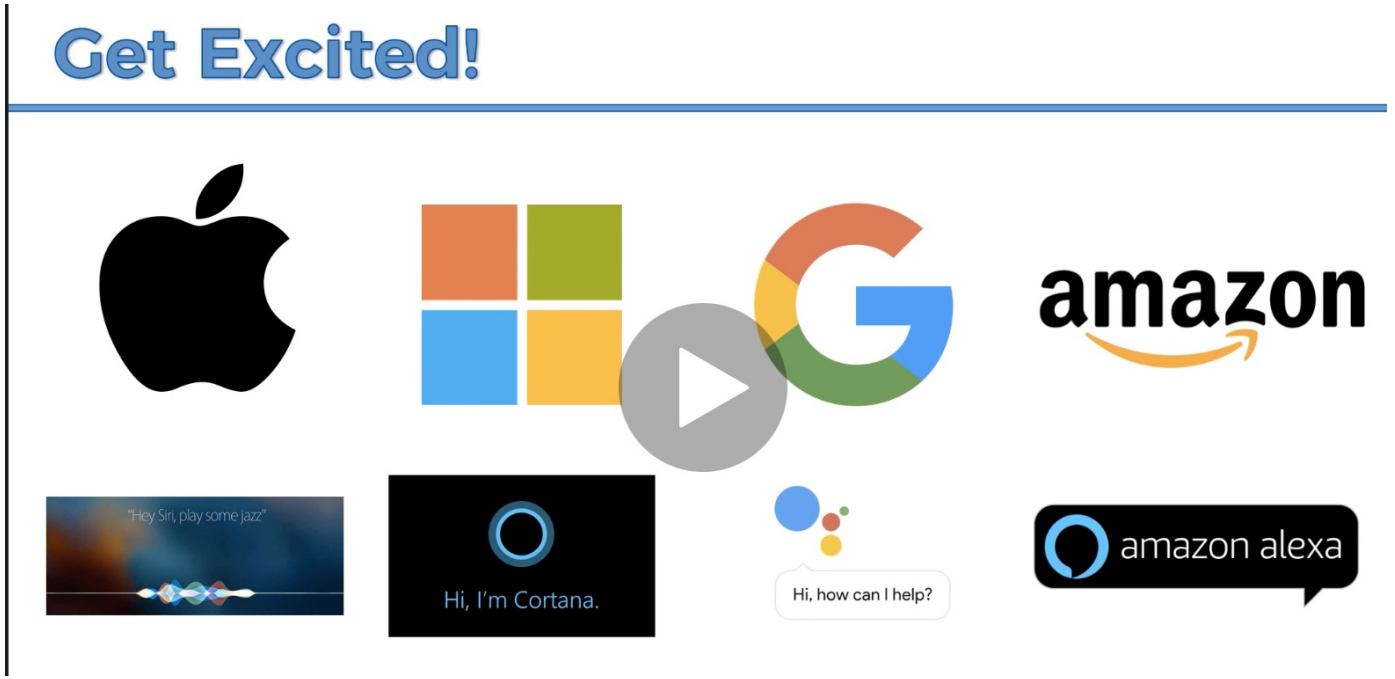
Deep Learning and NLP A-ZTM: How to create a ChatBot

Manikya Swathi Vallabhajosyula

***NOTE: Most of the images in the presentation were screenshots from the tutorial video*.**

Popular voice bots?

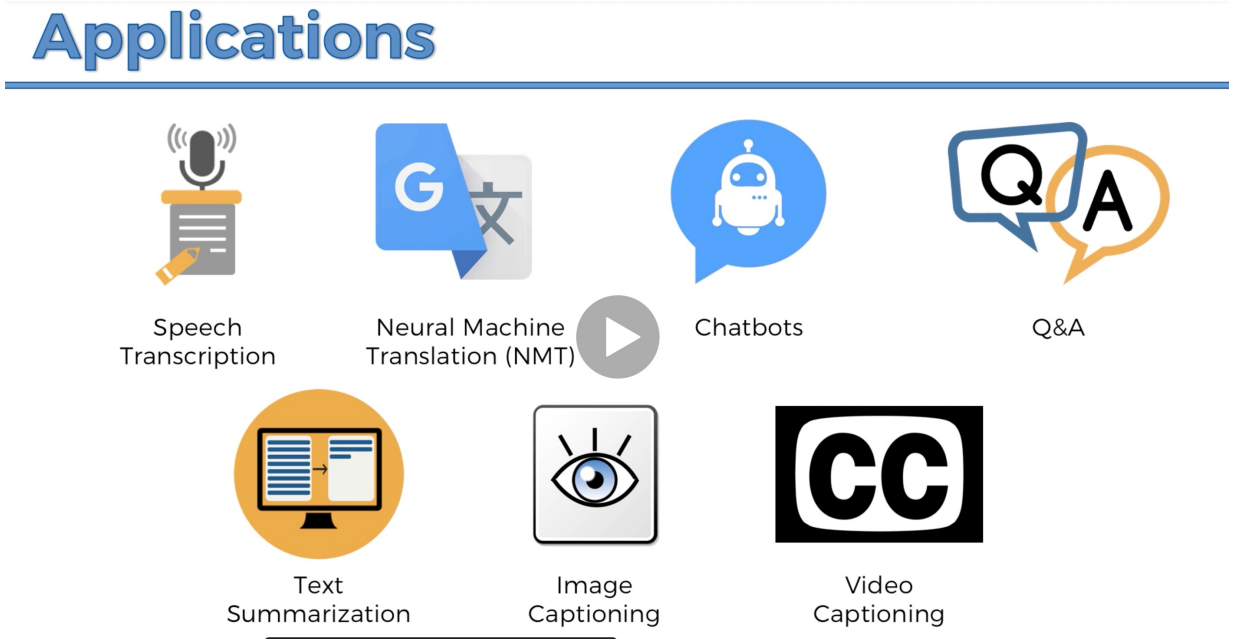
- Apple - Siri
- Google – Ok Google
- Amazon – Alexa
- Microsoft – Cortana



How? - Natural Language Processing (NLP)

Where is NLP used?

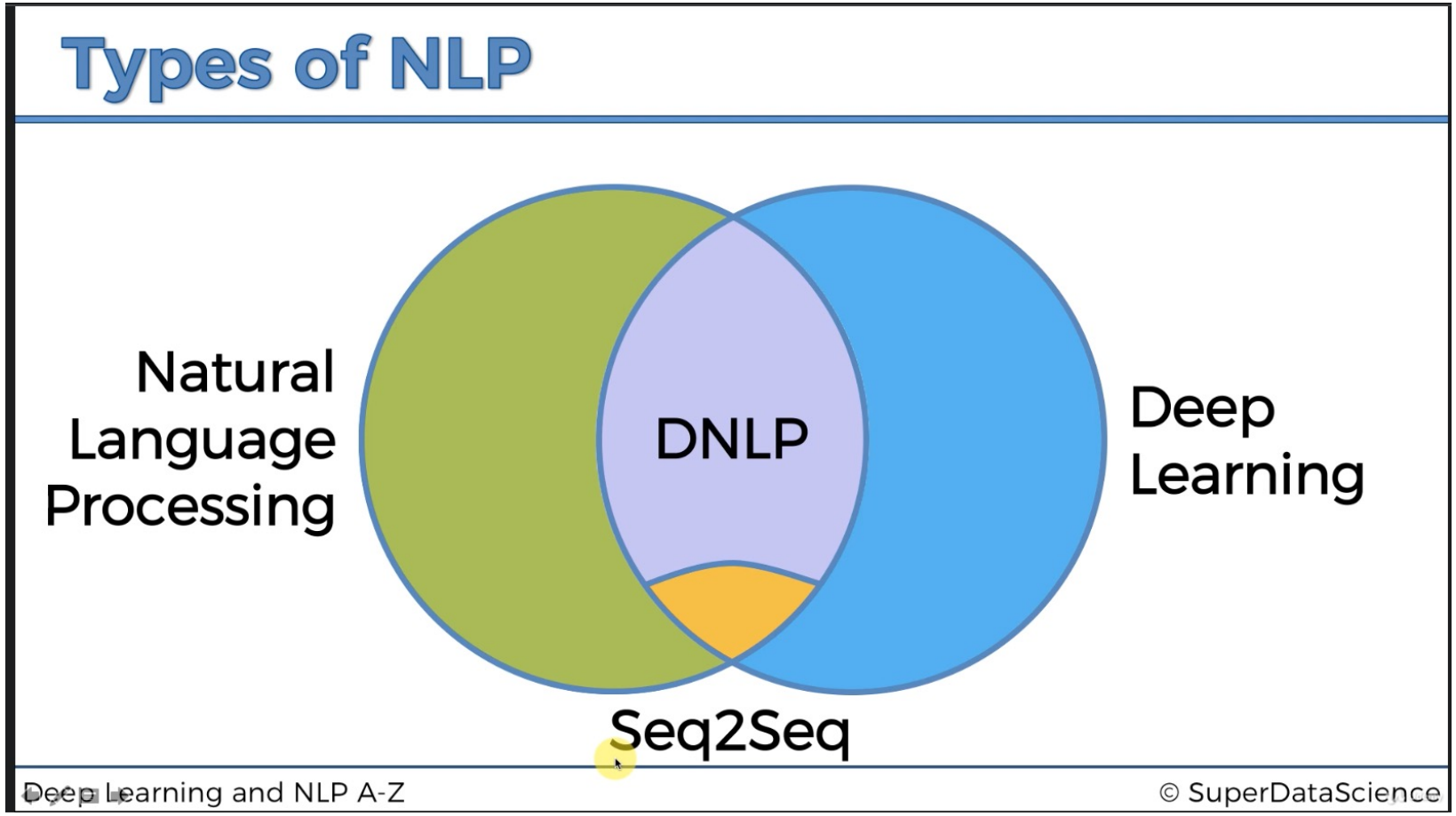
- Speech to text
  - Speech Transcriptions
  - Video captioning(combination)
- Text Processing
  - Neural Machine Translation
  - Chatbots
  - Q&A
  - Text summarization
- Video Processing
  - Image captioning
  - Video captioning(combination)



[BONUS] Free Material

- SDS 123: How to be Unstoppable: Data Science, Reckless Commitment & Artificial Intelligence
  - https://www.superdatascience.com/podcast/podcast-unstoppable-data-science-reckless-commitment-artificial-intelligence

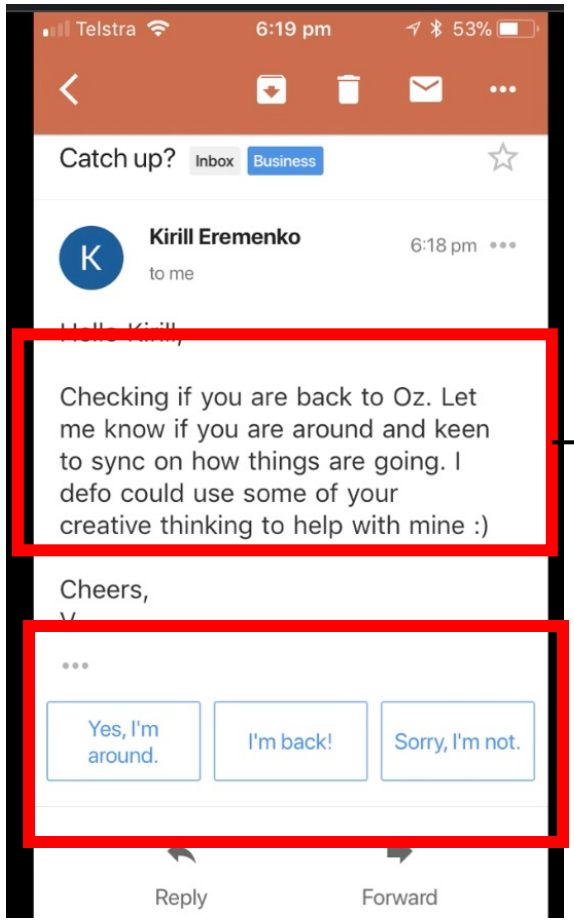
Deep NLP Intuition



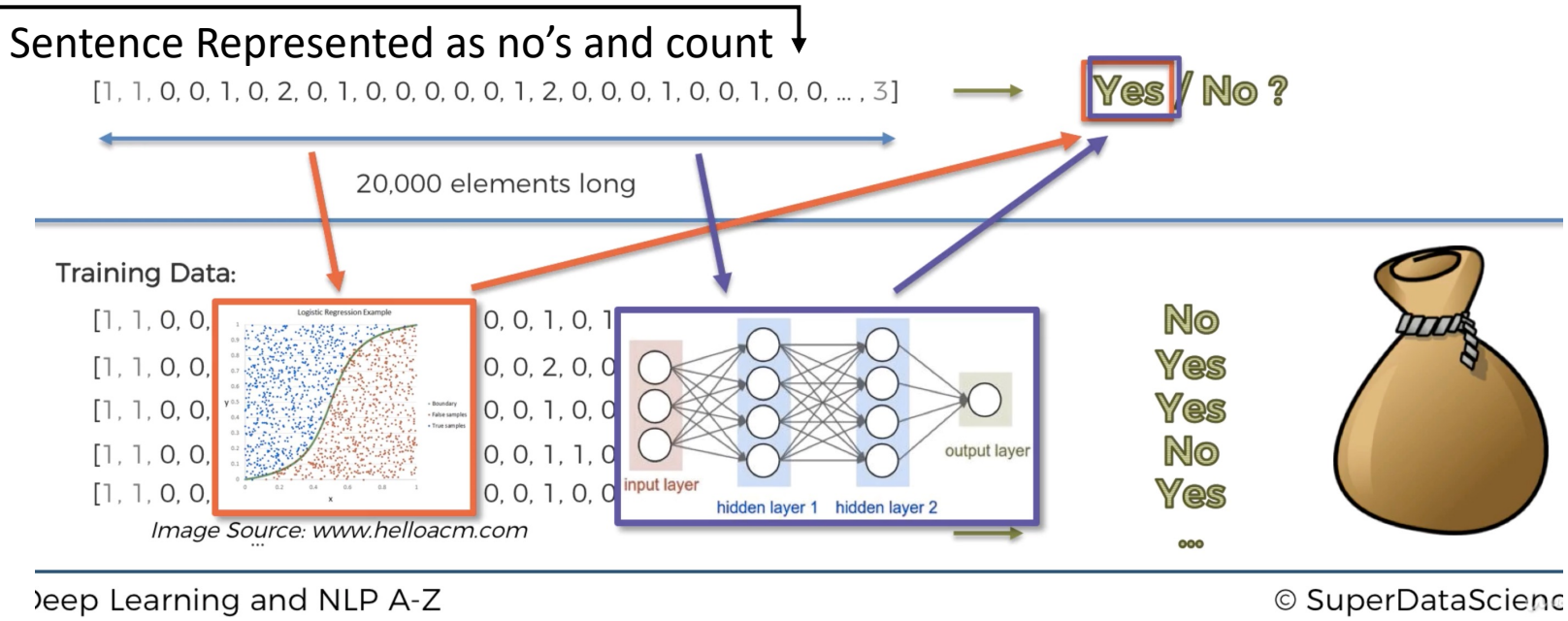
Classical VS Deep NLP- Models

- Classical
  - If/else blocks
  - Speech analysis by audio frequency model analysis
  - Back of words classification
- Deep Neural Models
  - Convolutional Neural Networks – classification
  - Sequence to Sequence Models – mostly text parsing and generation (end-to-end model)

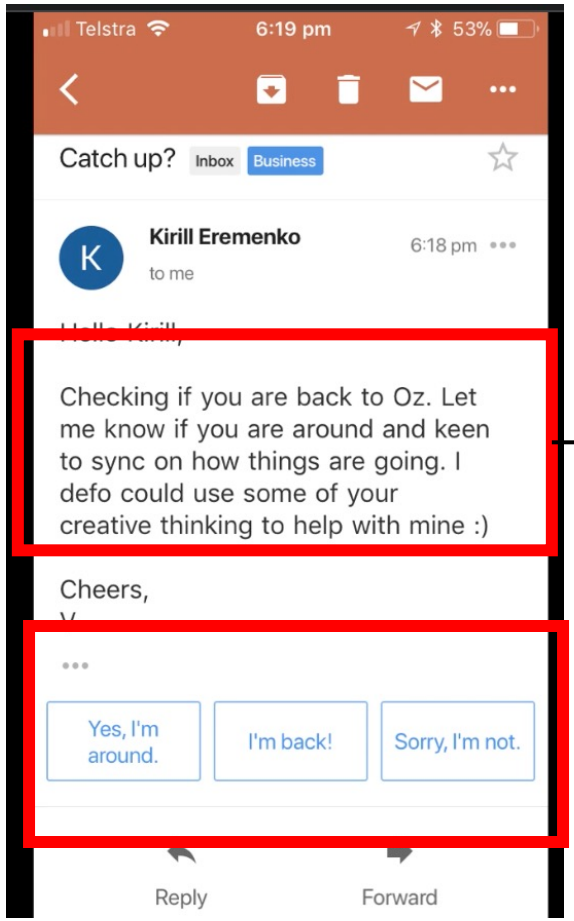
Text Generation – Bag-of-words: Generate yes/no



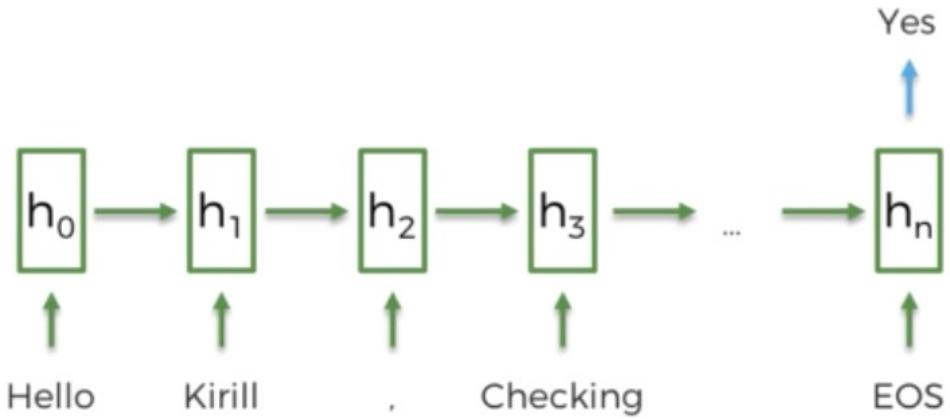
Bag-of-words: Looks at works surround the target words and computes scores



Text Generation – Deep-Neural Network Model

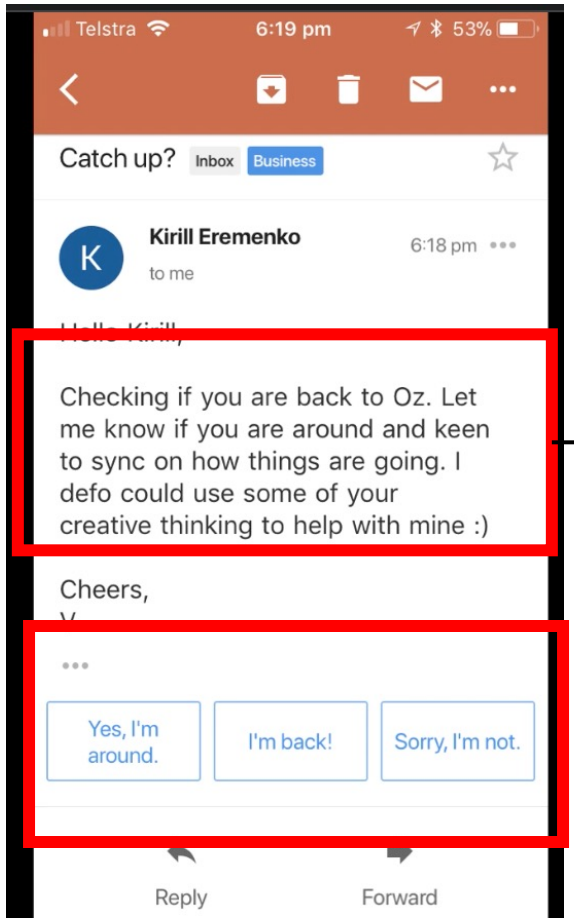


Seq-2-seq: Parses input sentence -> [Encoder] encodes it as H
- > [Decoder] decode the message H and produce ->**YES?NO**

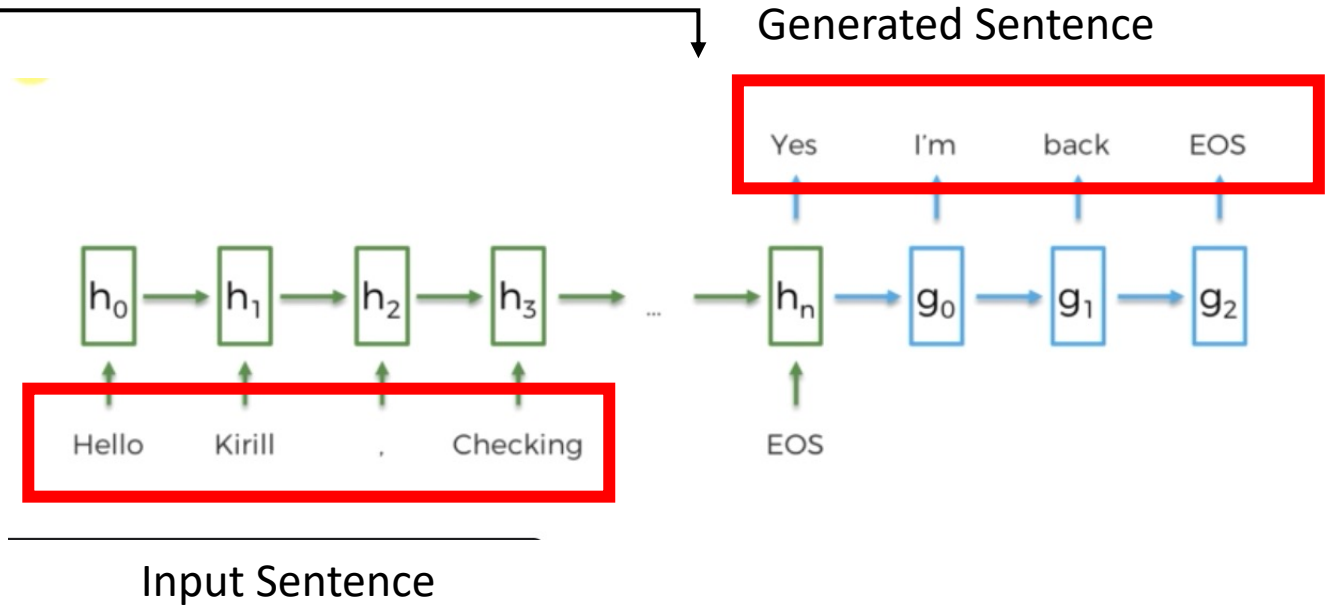


CAN DO BEYOND!!!

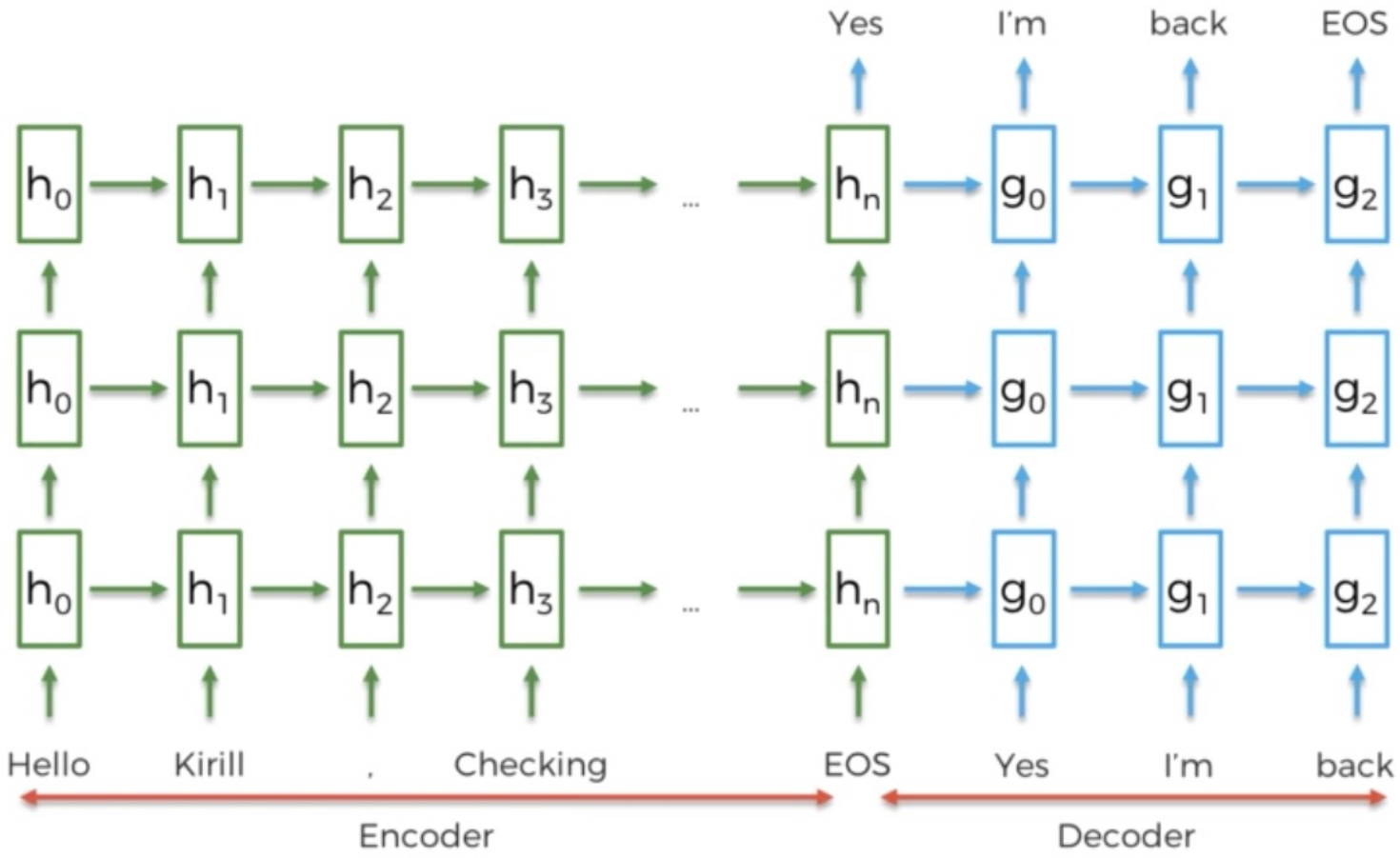
Text Generation – Bag-of-words: Generate yes/no



Seq-2-seq: Generate Full Response

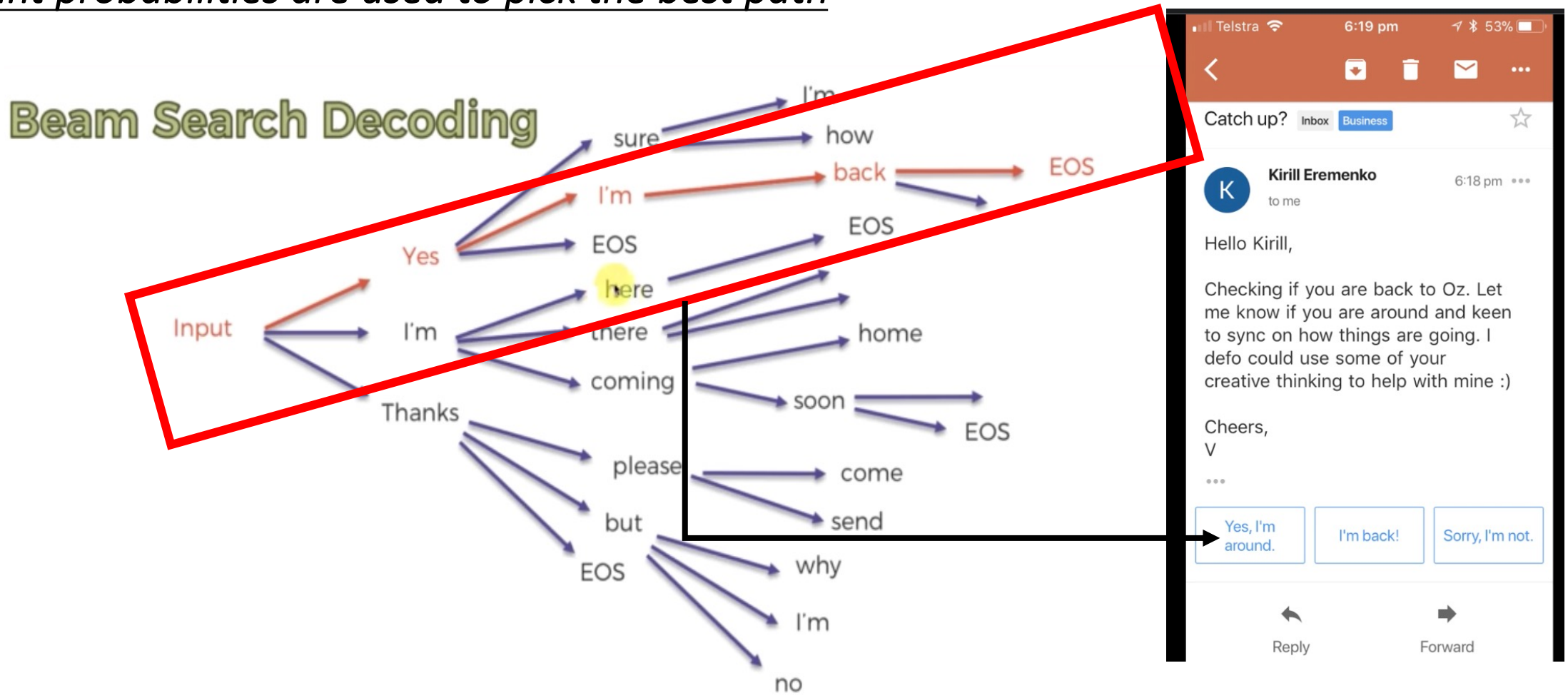


Can Go many layer deep and generate better self-learning models



Beam Search – popular decoding technique

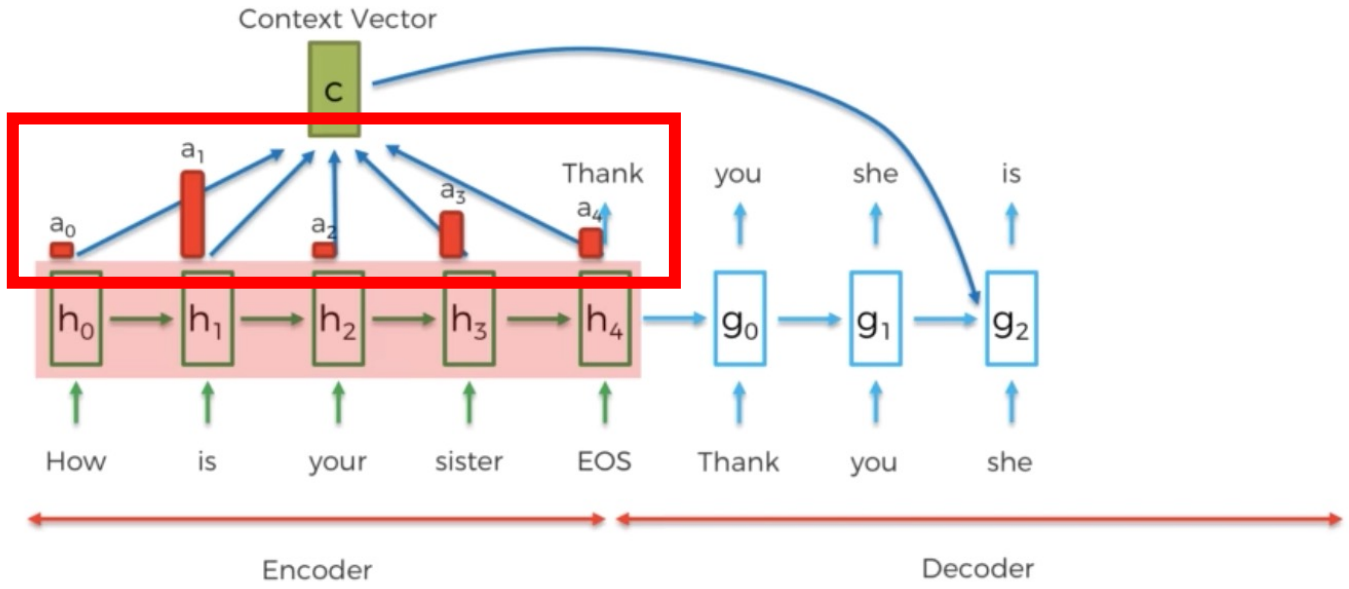
Joint probabilities are used to pick the best path



ATTENTION!!!

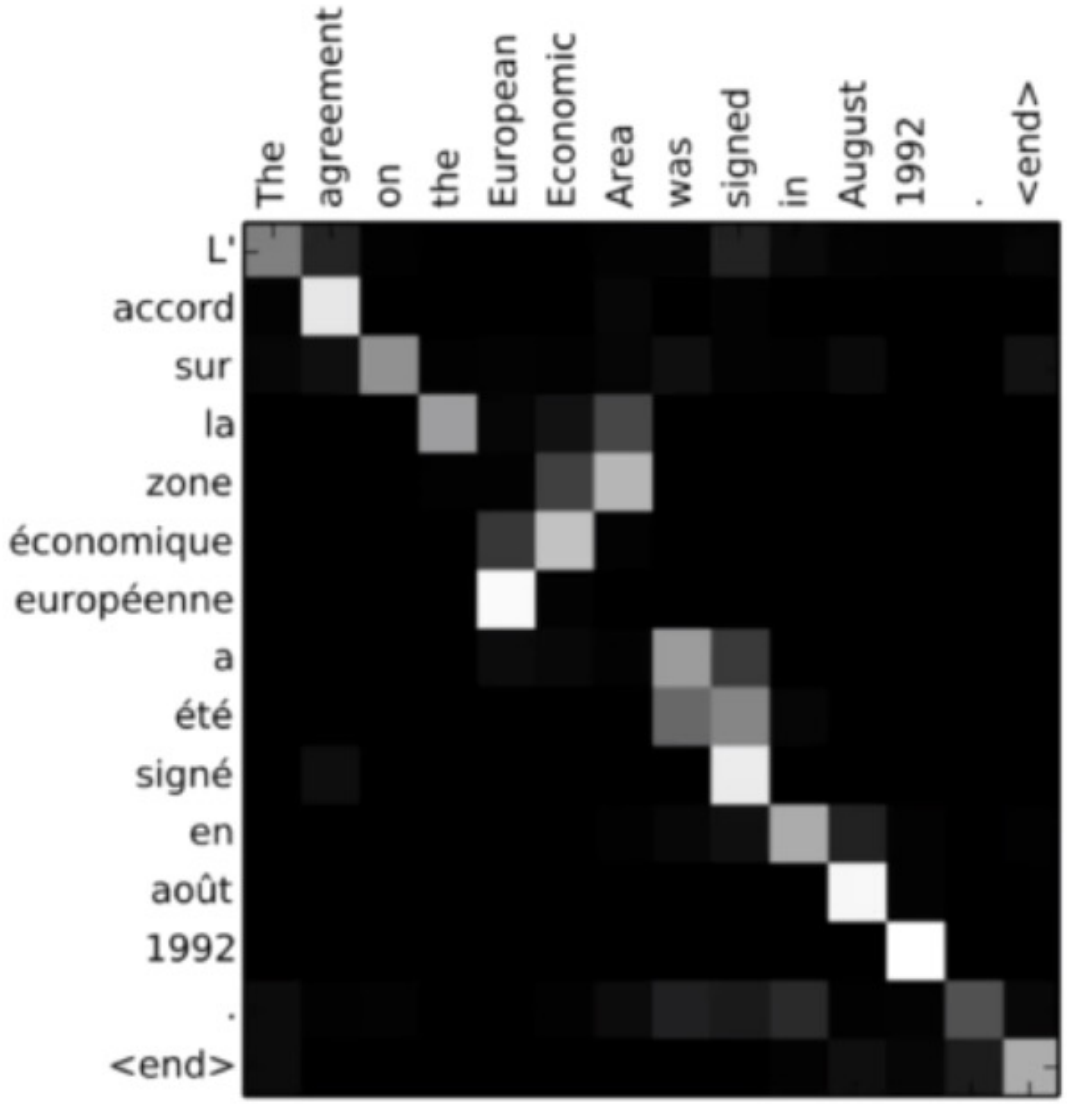
- Attention is all you need! – very very very popular idea proposed by google that revolutionized NLP + Seq-2- Seq models
  - https://arxiv.org/pdf/1706.03762.pdf

The red bars height indicates which words should be given more ***attention***!



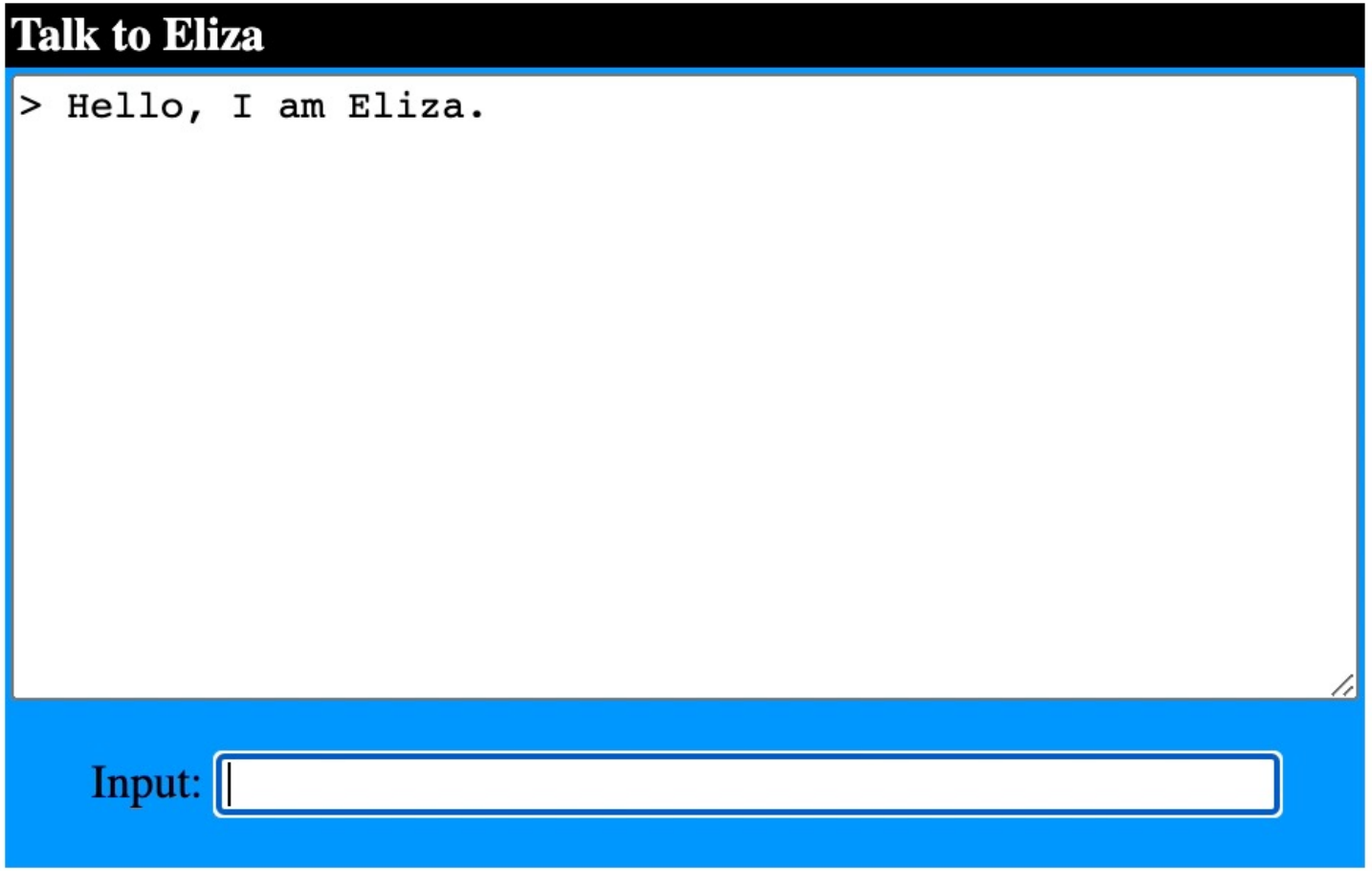
ATTENTION!!!

- Other application?? – Machine translation
  - E.g.: English to french (from https://arxiv.org/pdf/1706.03762.pdf)
  - Which word should have more weight for current word translations
  - Shows 1-to-1 association in translation



This is ELIZA the Rogerian Therapist (Rule-Based Chat Bot)

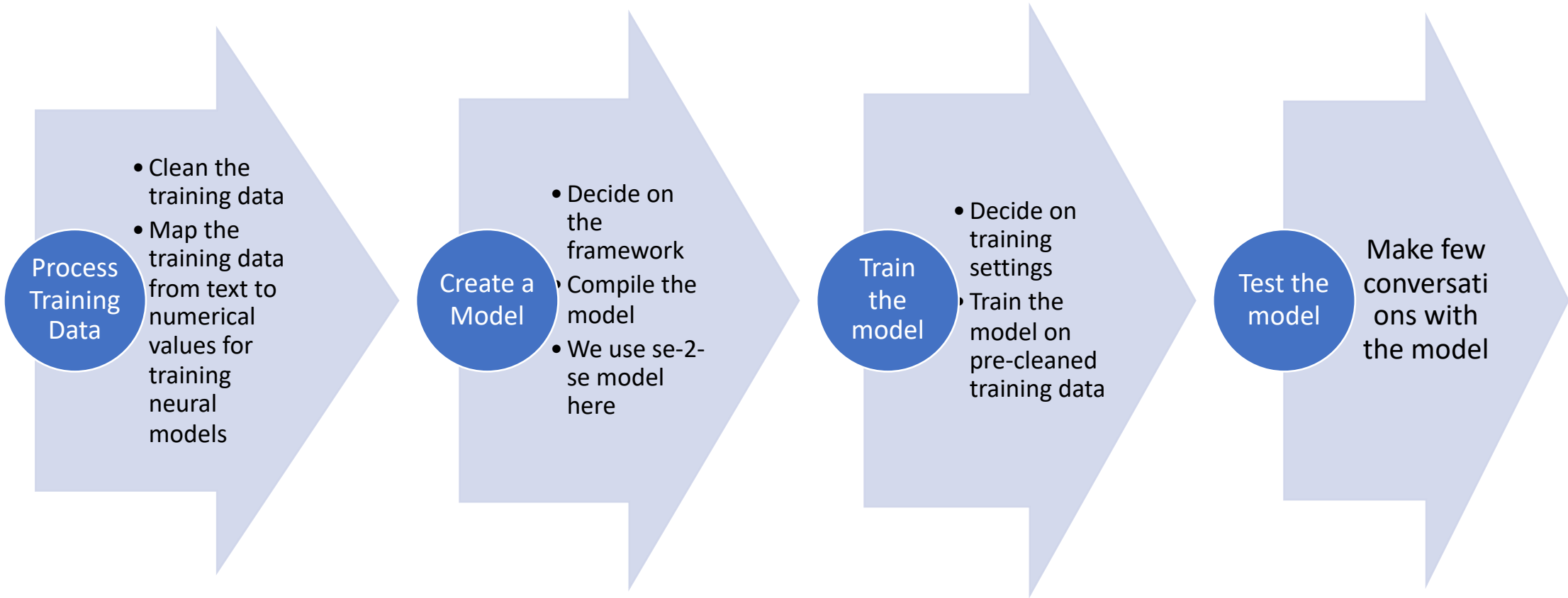
- http://psych.fullerton.edu/mbirnbaum/psych101/eliza.htm
- Play ELIZA.mov video from the submitted folder



5 mins chat with Eliza - Analysis

- We can see that it’s a simple rule mased machine that responds to certain keywords
  - When the use types in “sad”, “love”, ”happy”, ELIZA responds “How long are in in <word>?”
  - And when it doesn't have a specific rule, it responds “Tell me more..” (default)
  - It also stores few conversations, extracts phrases and asks ”what do you mean ‘i am sad’ these days”
- This is simple right! Now lets look at the framework that could be build to create such a bot…

Basic workflow for creating Chatbot



Steps to Set up a chat bot – Download data

- Install Anaconda
- Install Tensorflow in virtual environment
- Download data set -
  https://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html

"Cornell Movie--Dialogs Corpus"

- 220,579 conversational exchanges between 10,292 pairs of movie characters [WILL USE THIS]
- involves 9,035 characters from 617 movies
- in total 304,713 utterances
- movie metadata included: [NOT NEEDED]
  - genres
  - release year
  - IMDB rating
  - number of IMDB votes
  - IMDB rating
- character metadata included:
  - gender (for 3,774 characters)
  - position on movie credits (3,321 characters)
- see README.txt (included) for details

Sample training data x 1

```
L1045 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ They do not!
L1044 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ They do to!
L985 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ I hope so.
L984 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ She okay?
L925 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Let's go.
L924 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Wow
L872 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Okay -- you're gon
L871 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ No
L870 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ I'm kidding.  You                    ou just become this "persona"?  And you don't know how to quit?
L869 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Like my fear of we
L868 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ The "real you".
L867 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ What good stuff?
L866 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ I figured you'd get to the good stuff eventually.
L865 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Thank God!  If I had to hear one more story about your coiffure...
L864 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Me.  This endless ...blonde babble. I'm like, boring myself.
L863 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ What crap?
L862 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ do you listen to this crap?
L861 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ No...
L860 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Then Guillermo says, "If you go any lighter, you're gonna look like an extra on 90210."
L699 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ You always been this selfish?
L698 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ But
L697 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Then that's all you had to say.
L696 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Well, no...
L695 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ You never wanted to go out with 'me, did you?
L694 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ I was?
L693 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ I looked for you back at the party, but you always seemed to be "occupied".
L663 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Tons
L662 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Have fun tonight?
L578 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ I believe we share an art instructor
L577 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ You know Chastity?
L576 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Looks like things worked out tonight, huh?
L575 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Hi.
L407 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Who knows?  All I've ever heard her say is that she'd dip before dating a guy that smokes.
L406 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ So that's the kind of guy she likes? Pretty ones?
L405 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Lesbian?  No. I found a picture of Jared Leto in one of her drawers, so I'm pretty sure she's not
```

One pair –from a conversation between user "u0" and user "u2" u2 in movie "m0" – next we have character names (can ignore names) – then finally the "dialogues"

Sample training data x 2

“one full conversation” between user o0 and u2 in movie m0 referencing the lines from previous slieds

```
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L194', 'L195', 'L196', 'L197']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L204', 'L205', 'L206']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L271', 'L272', 'L273', 'L274', 'L275']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L280', 'L281']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L363', 'L364']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L365', 'L366']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L367', 'L368']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L401', 'L402', 'L403']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L404', 'L405', 'L406', 'L407']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L575', 'L576']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L577', 'L578']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L662', 'L663']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L693', 'L694', 'L695']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L696', 'L697', 'L698', 'L699']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L860', 'L861']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L862', 'L863', 'L864', 'L865']
u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L866', 'L867', 'L868', 'L869']
```

Requirements

- Install:
  - Conda, python 2.7
  - Create new environment with python 2.7
  - Load the new environment
  - Install Tenserflow1.0 (Apparently some modules in this example are stale for TF2.0 and code cannot be ported)
- Library dependency :
  - re (regular expressions for data cleaning)
  - time (to measure the performance)
  - numpy (for mathematical operations)
  - Tensorflow (use existing deep learning framework to train the chat bot)
- WE DEAL WITH THE CHAT BOT TRAINING AS Q-A Training

1. Process Training Data

- This step involves
  - 1. Extracting the question and answers from the conversation
    - Example:
    - Conversation: u0 +++$+++ u2 +++$+++ m0 +++$+++ ['L194', 'L195', 'L196', 'L197']
    - Dialogues in Conversation:

    L197 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Okay... then how 'bout we try out some French cuisine. Saturday? Night?

    L196 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Not the hacking and gagging and spitting part. Please.

    L195 +++$+++ u2 +++$+++ m0 +++$+++ CAMERON +++$+++ Well, I thought we'd start with pronunciation, if that's okay with you.

    L194 +++$+++ u0 +++$+++ m0 +++$+++ BIANCA +++$+++ Can we make this quick? Roxanne Korrine and Andrew B
  - Conversation as question answer:

One

Q: Can we make this quick? Roxanne Korrine and Andrew B
A: Well, I thought we'd start with pronunciation, if that's okay with you.

Two

Q: Well, I thought we'd start with pronunciation, if that's okay with you
A: Not the hacking and gagging and spitting part. Please..

Three

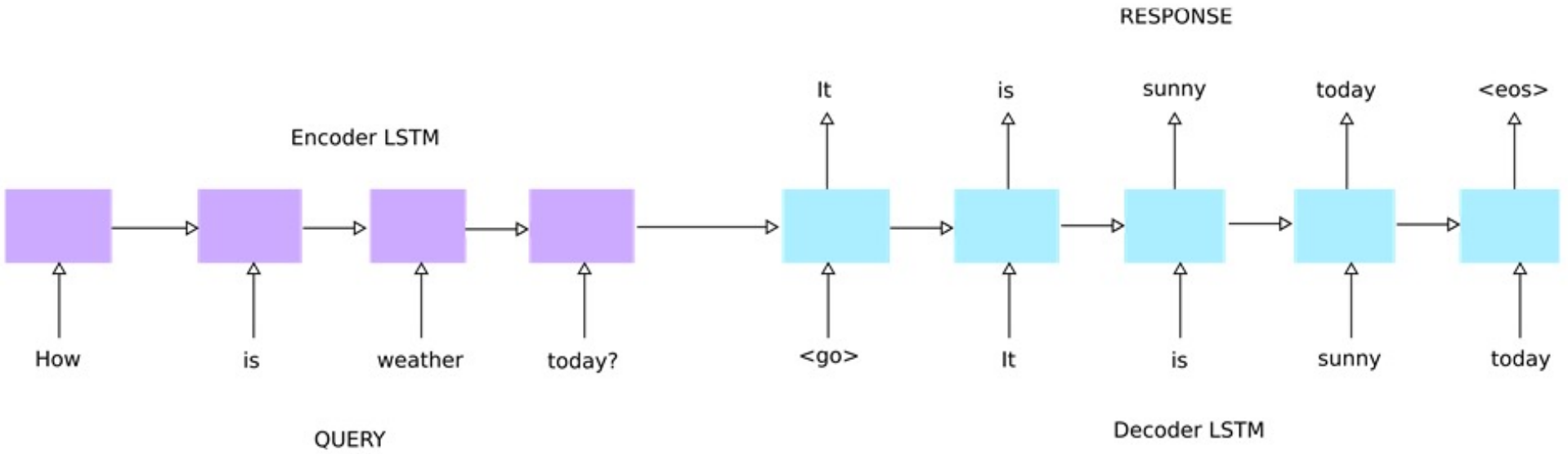
Q: Not the hacking and gagging and spitting part. Please..
A: Okay... then how 'bout we try out some French cuisine. Saturday? Night?

1. Process Training Data

- This step involves
  - 2. Identify all words in both questions and answers
  - 3. Delete words that have less than 5% occurrence in the training data
  - 4. Convert text data to numerical data so that a neural model could be built
    - Convert words into sparce number vectors
    - Create an embedding with respect to training data for the words

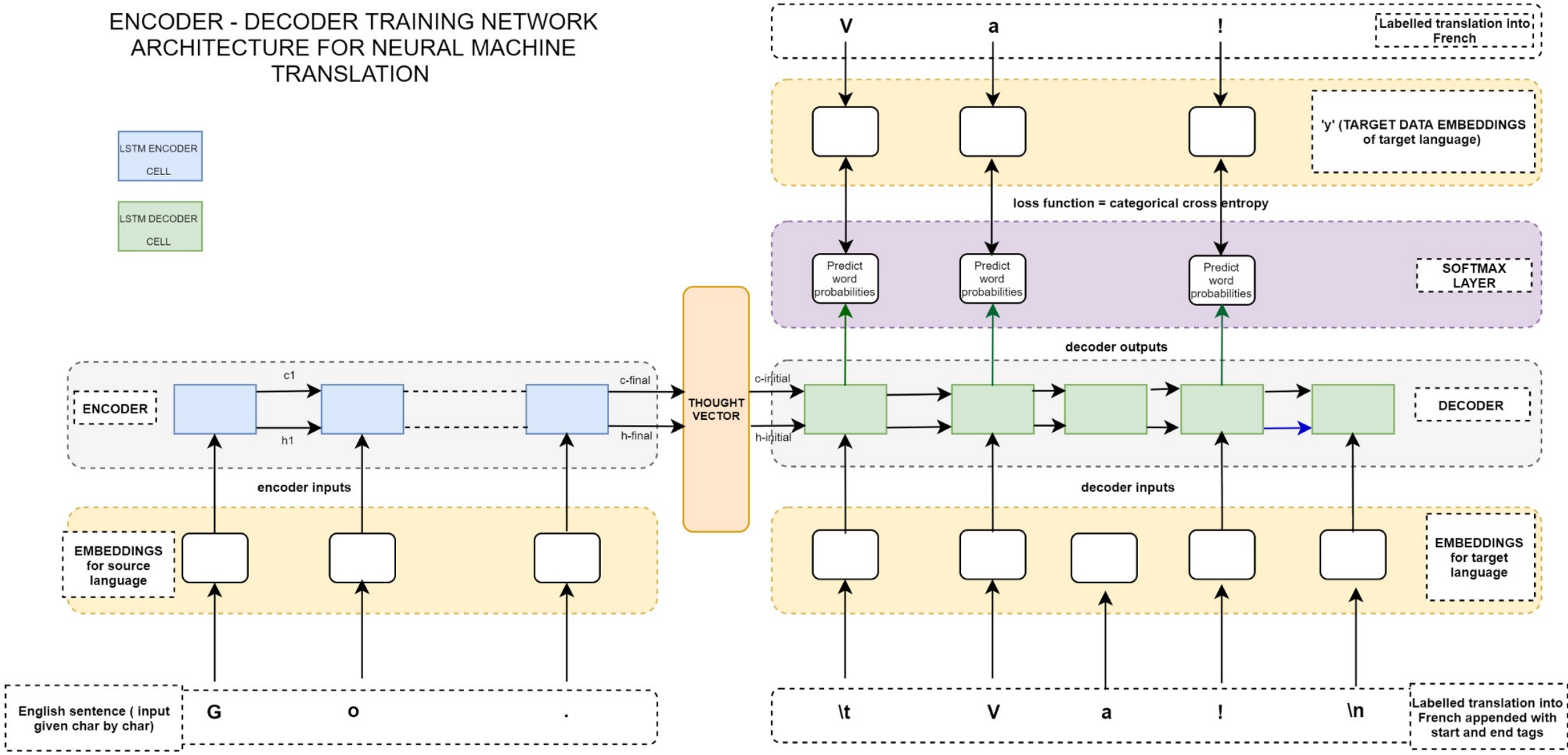
2. Create an encoder-decoder seq-2-seq model

- The following is the architecture of the model



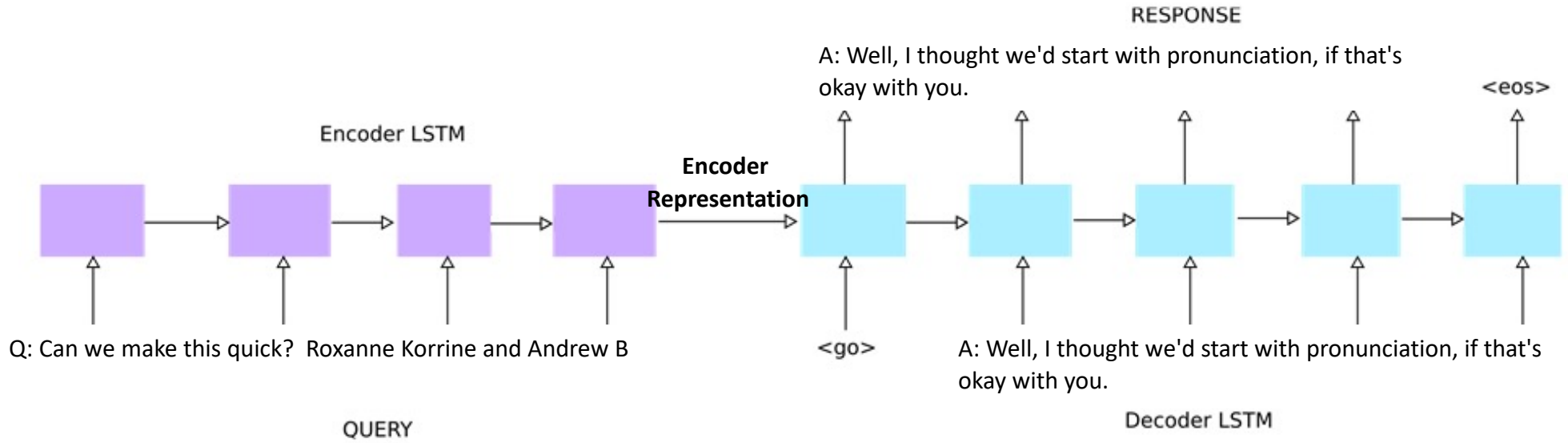
Detailed architecture is from here: https://towardsdatascience.com/how-to-implement-seq2seq-lstm-model-in-keras-shortcutnlp-6f355f3e5639

ENCODER - DECODER TRAINING NETWORK ARCHITECTURE FOR NEURAL MACHINE TRANSLATION



3. Training

- Feed the question answer pairs to the encoder – decoder model as



Takes input sequence and generates an encoder representation

Takes encoder representation and respective answer from training set to tune the model to generate the same target answer – learning to decode a question with respect to the answer

Training Complexities

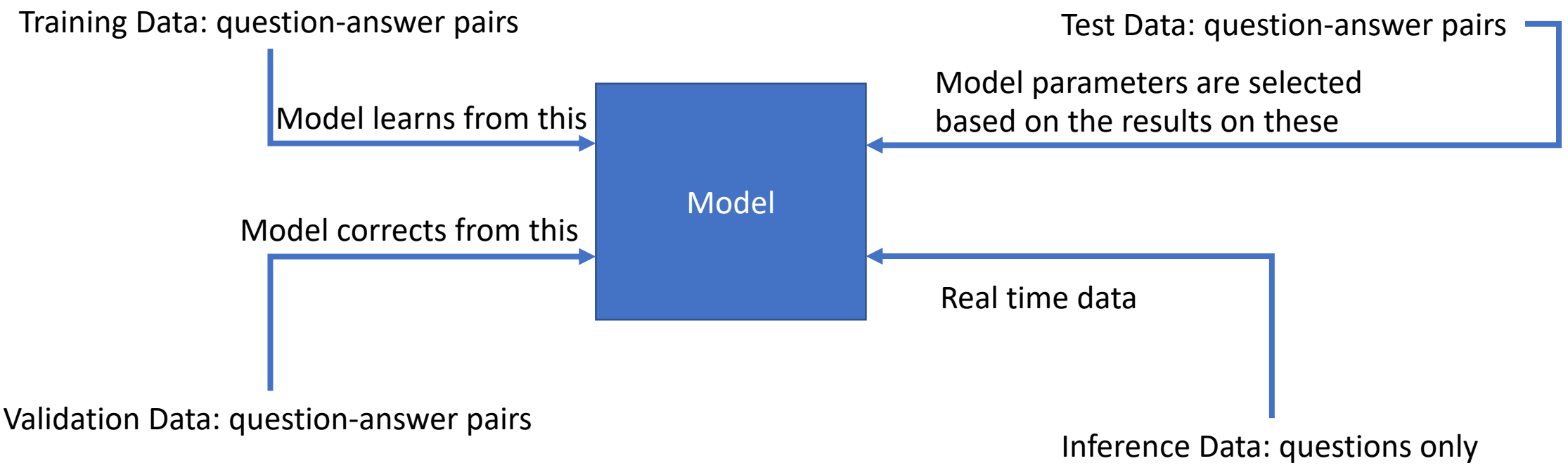
- How many examples should the model see to make a fair human-like answers that are relevant to questions? – size of training data
- How many times the identified training data should be read by the model that it learns about relevant conversations? – size of epochs (repetitive learning)
- How many sample should the model read and once to gain information (this is sort of – how much can it read at once to remember before going to next examples?) – Batch size
- How fast should the model learn the information? + How much of the information which produces wrong results should be forgotten? – learning rate + learning rate decay

So what are the best values – to get best results - ?

Multiple Training Variations

- Try several combinations of all the hyper-parameters: epochs, batches, learning-rate, learning-rate-delay, optimizer (softmax), training vs validation sample.
- Run tests with all sample models – the one with highest accuracy – final model
- Use that model for inference (i.e. for new unseen conversations)

Kinds of Data



Link to FUN PART

- The chat bot: https://colab.research.google.com/drive/1PqTO4z_uydanohwE6jcLLzkGUA27hYXp?usp=sharing
- NOTE: THIS TUTORIAL IS DONE IN PYTORCH AND NOT TESORFLOW
  - It is easier to run code on colab as it makes it easier for running it – no installing software is needed
  - TF1.0 from tutorial needs python2.7 and colab is on python3.5, hence had to shift to a pytorch example – But the frame work implement is same
- Video 1: How to setup and Run
- Video 2: Running another time with full data set
- Video 3: Full Data Set Results
- Video 4: Sort of cool conversations – with 1200 epochs
- Video Course Completion

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Fri Nov 12 9:51:21 PM
ChatBot.ipynb - Colaboratory
colab.research.google.com/drive/1PqTO4z...uydanohwE6jcLLzkGUA27hYXp?usp=sharing
Incognito
Update

ChatBot.ipynb

This code is re-factored from tutorial: https://medium.com/@ritidass29/create-your-chatbot-using-python-nltk-88809fa621d1

Colab: https://colab.research.google.com/github/pytorch/tutorials/blob/gh-pages/_downloads/chatbot_tutorial.ipynb#scrollTo=CPvHMPvSl9zy

```
[ ]  # JUST CHANGE THIS-> RE-RUN THE CODE
     n_iteration = 100
```

```
[ ]  from __future__ import absolute_import
     from __future__ import division
     from __future__ import print_function
     from __future__ import unicode_literals

     import torch
     from torch.jit import script, trace
     import torch.nn as nn
     from torch import optim
     import torch.nn.functional as F
     import csv
     import random
```

Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri Nov 12 9:58:40 PM

ChatBot.ipynb - Colaboratory

colab.research.google.com/drive/1PqTO4z_uydanohwE6jcLLzkGUA27hYXp?usp=sharing#scrollTo=GGlivA4G6SBA

Incognito Update

ChatBot.ipynb

File Edit View Insert Runtime Tools Help Last saved at 21:58

Comment Share

+ Code + Text

RAM Disk Editing

```
# Configure training/optimization
clip = 50.0
teacher_forcing_ratio = 1.0
learning_rate = 0.0001
decoder_learning_ratio = 5.0

print_every = 1
save_every = 500

# Ensure dropout layers are in train mode
encoder.train()
decoder.train()

# Initialize optimizers
print('Building optimizers ...')
encoder_optimizer = optim.Adam(encoder.parameters(), lr=learning_rate)
decoder_optimizer = optim.Adam(decoder.parameters(), lr=learning_rate * decoder_learning_ratio)
if loadFilename:
    encoder_optimizer.load_state_dict(encoder_optimizer_sd)
    decoder_optimizer.load_state_dict(decoder_optimizer_sd)

# Run training iterations
```

Chrome File Edit View History Bookmarks Profiles Tab Window Help Fri Nov 12 10:11:07 PM

ChatBot.ipynb - Colaboratory

colab.research.google.com/drive/1PqTO4z_uydanohwE6jcLLzkGUA27hYXp?usp=sharing#scrollTo=-2DQ__YBsb2X

Incognito Update

ChatBot.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

..
sample_data
formatted_movie_lines.txt
movie_conversations.txt
movie_lines.txt

Disk 66.10 GB available

+ Code + Text

RAM Disk

Editing

```
Iteration: 92; Percent complete: 92.0%; Average loss: 3.8819
Iteration: 93; Percent complete: 93.0%; Average loss: 3.7951
Iteration: 94; Percent complete: 94.0%; Average loss: 3.9551
Iteration: 95; Percent complete: 95.0%; Average loss: 3.6140
Iteration: 96; Percent complete: 96.0%; Average loss: 3.9067
Iteration: 97; Percent complete: 97.0%; Average loss: 3.3685
Iteration: 98; Percent complete: 98.0%; Average loss: 3.7116
Iteration: 99; Percent complete: 99.0%; Average loss: 3.9647
Iteration: 100; Percent complete: 100.0%; Average loss: 3.8551
```

[7] 2m

```
# Set dropout layers to eval mode
encoder.eval()
decoder.eval()

# Initialize search module
searcher = GreedySearchDecoder(encoder, decoder)

# Begin chatting (uncomment and run the following line to begin)
evaluateInput(encoder, decoder, searcher, voc)
```

>

Executing (1s) Cell > evaluateInput() > raw_input() > _input_request() > recv() > recv_multipart()

udemy | Deep Learning and NLP A-Z™: How to create a ChatBot

Get certificate | Share

Course content

- 73. ChatBot - Step 42: Introduction to a new model & setup — 3min — Resources
- 74. ChatBot - Step 43: Chatbot model discussion — 8min
- 75. ChatBot - Step 44: Tensorboard — 4min
- 76. ChatBot - Step 45: Run the new chatbot model — 4min — Resources

Section 9: Other ChatBot Implementations
5 / 5 | 19min

Section 10: Annex 1: Artificial Neural Networks
8 / 8 | 1hr 18min

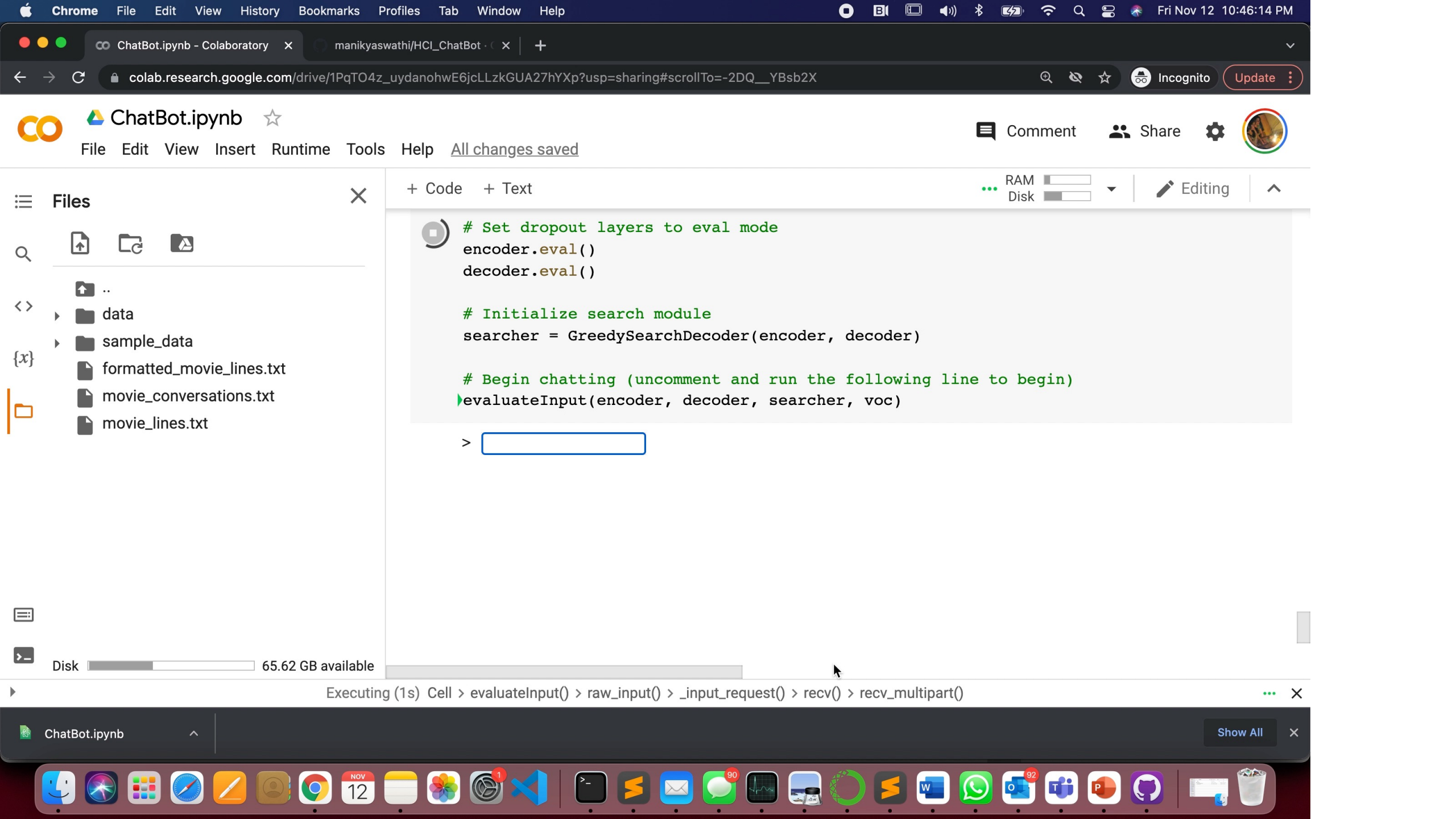
Section 11: Annex 2: Recurrent Neural Networks
6 / 6 | 1hr 12min

Overview | Q&A | Notes | Announcements

About this course

Learn the Theory and How to implement state of the art Deep Natural Language Processing models in Tensorflow and Python

Waiting for gtm.udemy.com...


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Fri Nov 12 10:46:14 PM
ChatBot.ipynb - Colaboratory
manikyaswathi/HCI_ChatBot
colab.research.google.com/drive/1PqTO4z_uydanohwE6jcLLzkGUA27hYXp?usp=sharing#scrollTo=-2DQ__YBsb2X
Incognito
Update
ChatBot.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
..
data
sample_data
formatted_movie_lines.txt
movie_conversations.txt
movie_lines.txt
Disk 65.62 GB available
+ Code + Text
RAM
Disk
Editing
# Set dropout layers to eval mode
encoder.eval()
decoder.eval()
# Initialize search module
searcher = GreedySearchDecoder(encoder, decoder)
# Begin chatting (uncomment and run the following line to begin)
evaluateInput(encoder, decoder, searcher, voc)
>
Executing (1s) Cell > evaluateInput() > raw_input() > _input_request() > recv() > recv_multipart()
ChatBot.ipynb
Show All

Link to FUN PART

- The chat bot: https://colab.research.google.com/drive/1PqTO4z_uydanohwE6jcLLzkGUA27hYXp?usp=sharing
- If links in previous slides did not play well, then try playing them from submitted folder
- Video 1: How to setup and Run (Play ChatBot1.mov video from the submitted folder)
- Video 2: Running another time with full data set (Play ChatBot2.mov video from the submitted folder)
- Video 3: Full Data Set Results (Play CourseCompletion.mov video from the submitted folder)
- Video 3: Sort of cool conversations – with 1200 epochs (Play ChatBot4.mov video from the submitted folder)
- Video Course Completion Play CourseCompletion.mov video from the submitted folder)

udemy

Certificate no: UC-bc8ed707-4eb1-47a2-a8dc-d387b4f02e40
Certificate url: ude.my/UC-bc8ed707-4eb1-47a2-a8dc-d387b4f02e40
Reference Number: 0004

CERTIFICATE OF COMPLETION

Deep Learning and NLP A-Z™: How to create a ChatBot

Instructors **Hadelin de Ponteves, Kirill Eremenko, Ligency | Team, Zillion Hand Team**

Manikya Swathi Vallabhajosyula

Date **Nov. 12, 2021**
Length **12 total hours**

Files Shared as ZIP

- ChatBot.ipynb - The colab code downloaded version
- TrainFile – Full Dataset folder
- TestFile – Sample Dataset folder
- All Videos:
  - ChatBot3.mov
  - ChatBot1.mov
  - CourseCompletion.mov
  - ChatBot2.mov
- GIT Repo: https://github.com/manikyaswathi/HCI_ChatBot.git