

Speech and sentiment analysis with Python APIs

Justin Shenk
@justinshenk

Data Science Meetup Münster, February 2018

About me

- Previously studied neuroscience (USA)
- Master thesis on deep learning with Peltarion (Stockholm)
at Osnabrück University
- Intel Software Innovator



Overview

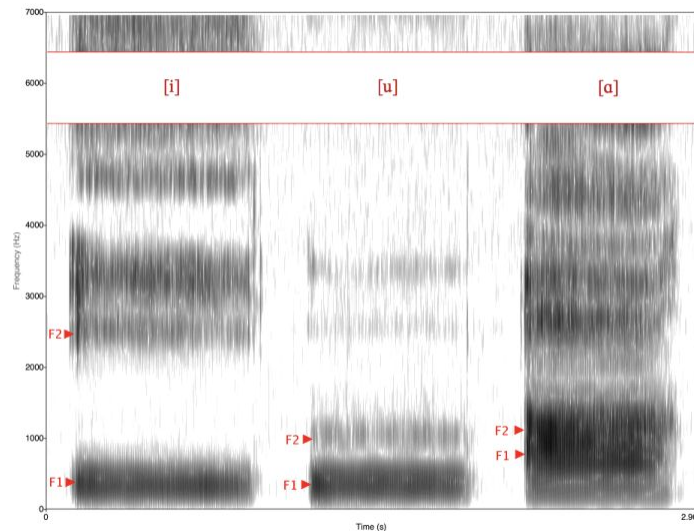
- Motivation
- Speech to Text
- Text Analysis
- Deep Learning / LSTM
- Demo
- Conclusion

Motivation

Use cases for personal (or corporate) speech recognition and text analysis

Speech Recognition - History

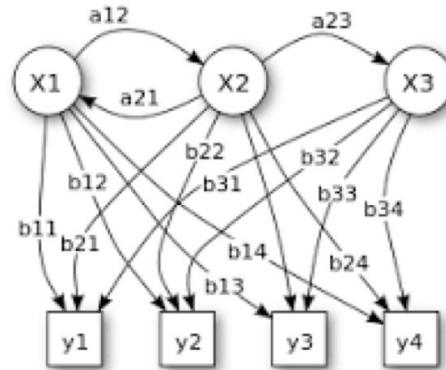
- 1950s – 10 words, Formants
- 1980s – 20,000 words, HMMs
- 2000s – Full Vocab – high error rate
HMM + Neural Net
- Modern – Full Vocab – low error rate
LSTM



Speech Recognition - Open Source

CMU Sphinx

- “Phone” based
- Probabilistic
- Runs Locally



Hidden Markov Chain Model - assigns label to each unit in sequence based on probabilistic approach



Speech Recognition - Using Sphinx

Input from audio .wav file

```
# use the audio file as the audio source
r = sr.Recognizer()
with sr.AudioFile(AUDIO_FILE) as source:
    audio = r.record(source)
```

Input from microphone

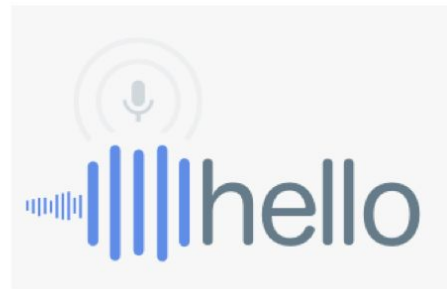
```
# obtain audio from the microphone
r = sr.Recognizer()
with sr.Microphone() as source:
    print("Say something!")
    audio = r.listen(source)
```

Call Sphinx function

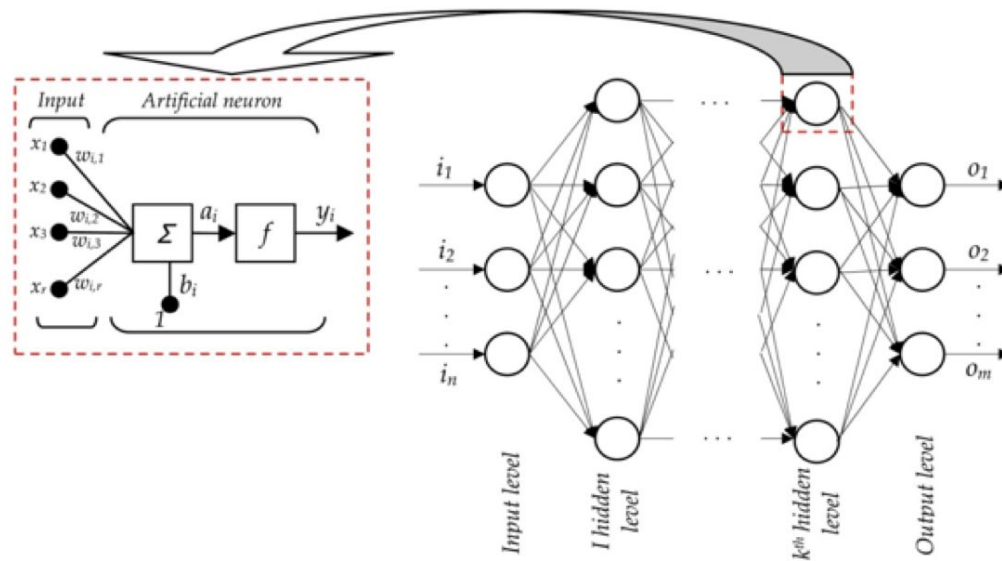
```
# speech recognition using Sphinx
speech_to_text = r.recognize_sphinx(audio)
```

Google Cloud Speech API

- Neural Network
- 110 Languages
- Restful API
- Google Cloud SDK



Deep Learning / Neural Networks





Using Cloud Speech API

- Console Frontend
- HTTP Request
- Google Cloud SDK

Convert your speech to text right now

Select a language and click "Start Now" to begin recording

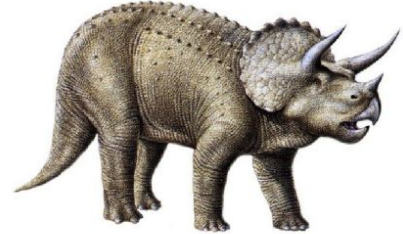
English (United States) 

 **START NOW**

```
r = sr.Recognizer()
with sr.Microphone() as source:
    audio = r.listen(source)
print("Google Speech Recognition thinks you said '" + r.recognize_google(audio) + "'")
```

Sentiment Analysis and Content Classification

Text Analytics

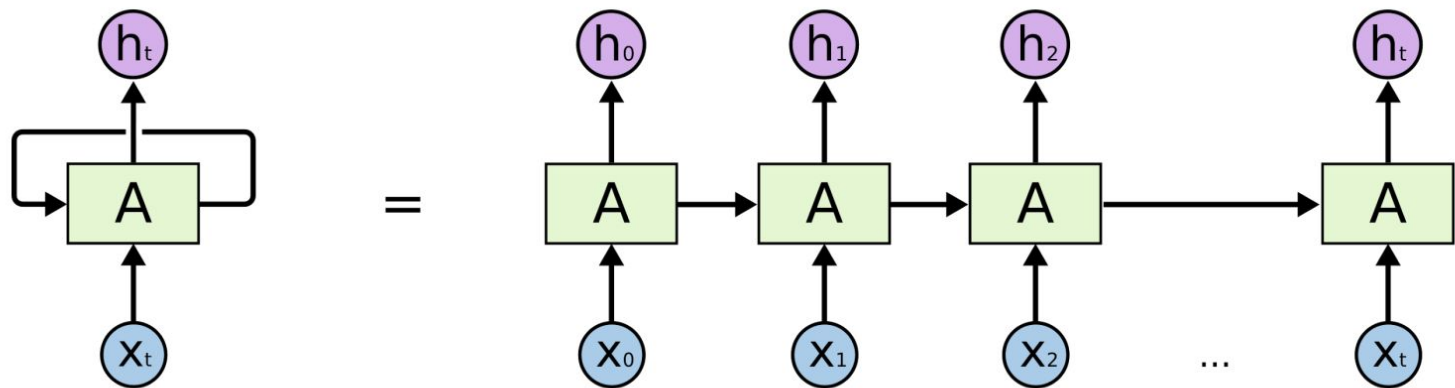


Neural Networks - LSTM

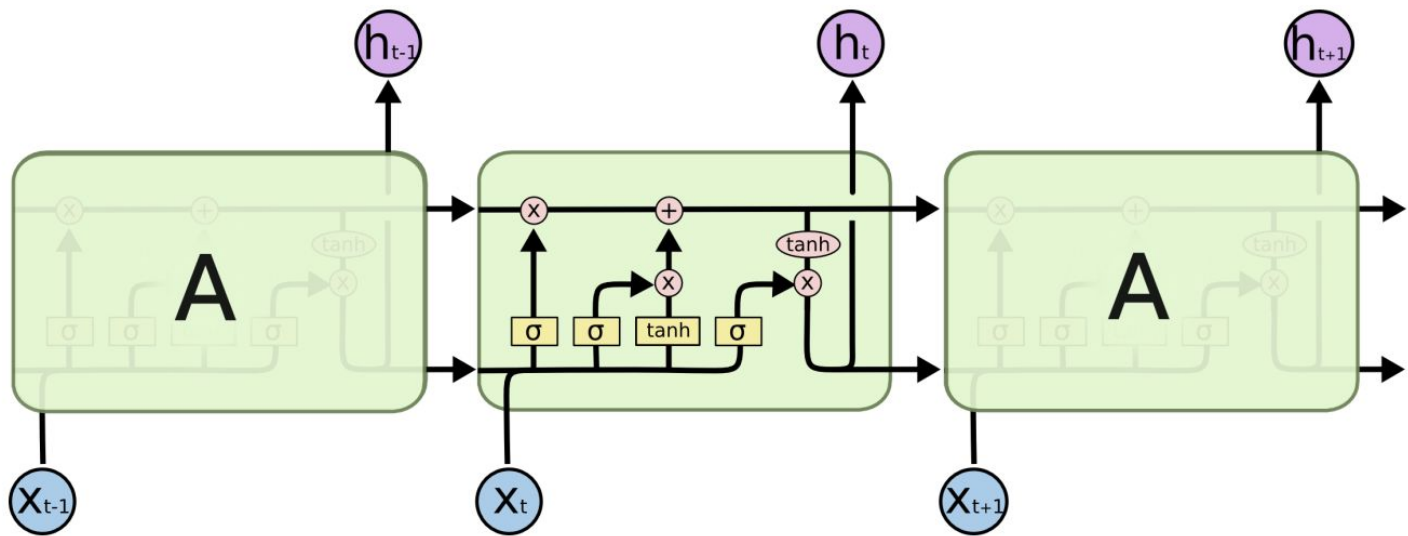
- Long Short Term Memory
- Type of RNN
- Contextual Analysis
- “Non volatile memory”

“I went to **France**... I learned how to
speak french”

Recurrent Neural Networks



LSTM Specific Structure



Result of LSTM

“Google, headquartered in Mountain View, unveiled the new Android phone at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.”

/Computers & Electronics

Confidence: 0.61

/Internet & Telecom/Mobile & Wireless

Confidence: 0.53

/News

Confidence: 0.53

Demo

Code: github.com/justinshenk/data-science-meetup-muenster



Record Audio



GCP ML APIs

- Speech to Text
- Sentiment & Topic Analysis



- Speech to Text
- Connect to GCP
- Receive Results

Results

- Speech recognition: Sphinx, Google Cloud Speech API
- Text Analysis: Google natural language API
- Audio input: An IoT gateway is an intermediate device between sensors and devices and the applications that create value from their data and access.

```
(env) priyanka@priyanka-desktop:~/gcp$ python speech_sphinx.py
ALSA lib pcm_dsnoop.c:686:(snd_pcm_dsnoop_open) unable to open slave
ALSA lib pcm_dmix.c:1029:(snd_pcm_dmix_open) unable to open slave
ALSA lib pcm.c:2266:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.rear
ALSA lib pcm.c:2266:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.center_lfe
ALSA lib pcm.c:2266:(snd_pcm_open_noupdate) Unknown PCM cards.pcm.side
ALSA lib pcm_dmix.c:1029:(snd_pcm_dmix_open) unable to open slave
Please wait. Calibrating microphone...
Say something!

Speech Recognition:
Sphinx thinks you said 'in a coyote gateways intermediate the fight between centers and devices in the application to create value from that that let's'
Time required for Sphinx analysis: 3.96572113037 sec
Google Speech Recognition thinks you said 'an IoT Gateway is an intermediate device between sensors and devices and the applications that create value from their data and access'
Time required for Google Speech Recognition: 3.41175913811 sec

Google ML API for Natural Language Processing:
magnitude: 0.10000000149
score: -0.10000000149

name          : /Computers & Electronics
confidence    : 0.870000004768
=====
(env) priyanka@priyanka-desktop:~/gcp$
```

“It’s more profound than, I don’t know,
electricity or fire.”

- Sundai Pichai, Google CEO

Questions?

Contact:

[linkedin.com/in/justinshenk](https://www.linkedin.com/in/justinshenk)
shenk.justin@gmail.com

