NUTS NARS and Speech

NUTS: raNdom dimensionality redUction non axiomaTic reasoning few Shot learner for perception

Dwane van der Sluis June 2023

Code & Supplementary Material: https://github.com/dwanev/NUTS.git

Motivation

- Children learn new words with just a few examples
- ONA (a software platform) was built on top of the assumption that it will be operating with insufficient knowledge and resources (AIKR).
- Can Open Non Axiomatic Reasoning for Applications (ONA) be used for perception?
- Can it be made sample efficient?

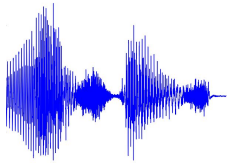
Assumed Knowledge

- That the audience knows about Open NARS for applications
  - If not see https://github.com/opennars/OpenNARS-for-Applications

- The the audience knows a little Narsese
  - If not see https://cis.temple.edu/~pwang/NARS-Intro.html

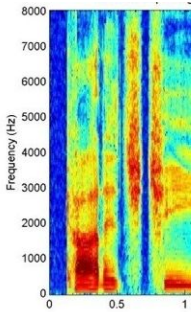
Sound Recap

Audio



16k
samples/sec

MEL
Spectrogram



8k
samples/sec

Dataset

List of words in the Speech Commands dataset v2:

’bed’, ’cat’, ’down’, ’five’, ’forward’, ’go’, ’house’, ’left’, ’marvin’, ’no’, ’on’, ’seven’, ’six’, ’tree’, ’up’, ’visual’, ’yes’, ’backward’, ’bird’, ’dog’, ’eight’, ’follow’, ’four’, ’happy’, ’learn’, ’nine’, ’off’, ’one’, ’right’, ’sheila’, ’stop’, ’three’, ’two’, ’wow’, ’zero’

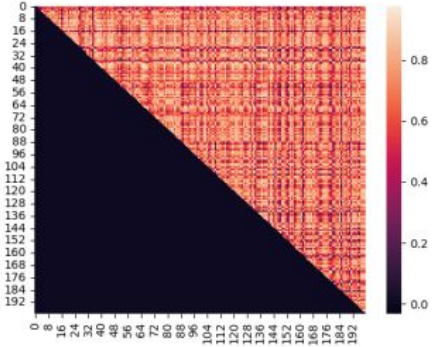
Consists of 105,829 utterances of 35 words, of between 1500 and 4000 utterance each, multiple speakers.

Warden, P.: Speech commands: A dataset for limited-vocabulary speech recognition. arXiv preprint arXiv:1804.03209 (2018) https://doi.org/10.48550/arXiv.1804.03209

Dataset

Audio Data is less self similar in the

Time frequency domain than we would expect.



(b) Correlation between each utterance mel values for 200 randomly chosen recordings of 'One' from the Standard Command dataset v2.

High Level Process

Build background knowledge inside ONA, *n* examples of each class.

Present ONA with a new unlabeled instance

Query ONA as to what the unlabeled instance is

Pseudo Code

Generate a random 8000x**D** matrix **R**

For Unlabelled class in class list:

Pick **1** random instance from the unlabeled class **U**

Pick ***n*** random samples from the class

Convert the ***n*** samples into Narsese by

Load waveform, convert into Mel Spectrogram **S** # dimension [8000,1]

Reduced Dimension Matrix **E** = **S** * **R** # dimension [**D**,1]

E = normalised(**E**) # normalises all values to between 0.0 and 1.0

Convert **E** -> Narsese Statements # shown on next slide

Add Instance Label Narsese statement (“isa”) for all the labeled samples

Narsese Statements - > Nalifier -> Filtered Narsese Statements -> ONA

Convert **U** into Narsese as above (but with no “isa” declaration)

Query ONA with < {**U**} -> Label_1>?

Encoding into Narsese

< {Word_One_Utterance_A} -> [p1]>. %0.9%

< {Word_One_Utterance_A} -> [p2]>. %0.51%

< {Word_One_Utterance_A} -> [NOTp3]>. %0.8% 1 Narsese statement for each property p1, p2 … p**D**

……

< {Word_One_Utterance_B} -> [p1]>. %0.9%

…….

< {Word_Two_Utterance_C} -> [p1]>. %0.9%

…….

< {Word_One_Utterance_A} -> Label_ONE >. State what A, B are

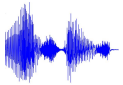
< {Word_One_Utterance_B} -> Label_ONE >.

……

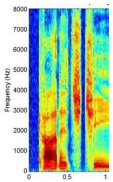
< {Unlabbeld_Utterance_X} -> Label_1>? Query What C is similar to (for every class)

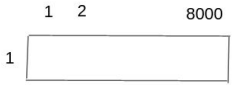
Consider correct if highest > 0.5 and is the correct class

For Narsese see https://cis.temple.edu/~pwang/NARS-Intro.html

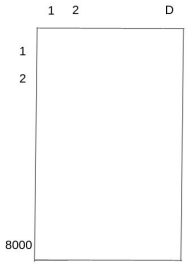


Audio

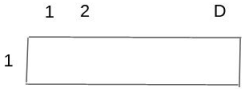




MEL
Spectrogram
1x8000



Random
8000xD



Reduced
Embedding
1xD

Narsese

Preprocess

Nalifier.py
(part of ONA)

ONA

Is A? not What Is?

In a 35 class example, we asked ONAs

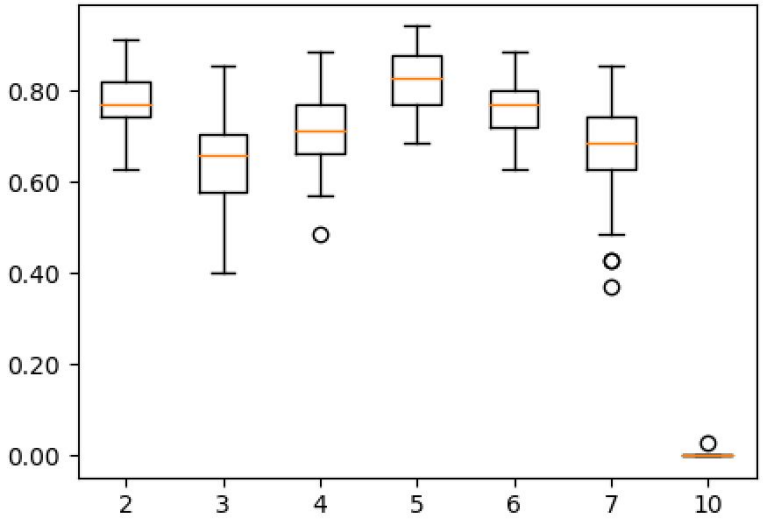
“is this a X?”,

Rather than

“what is this?”

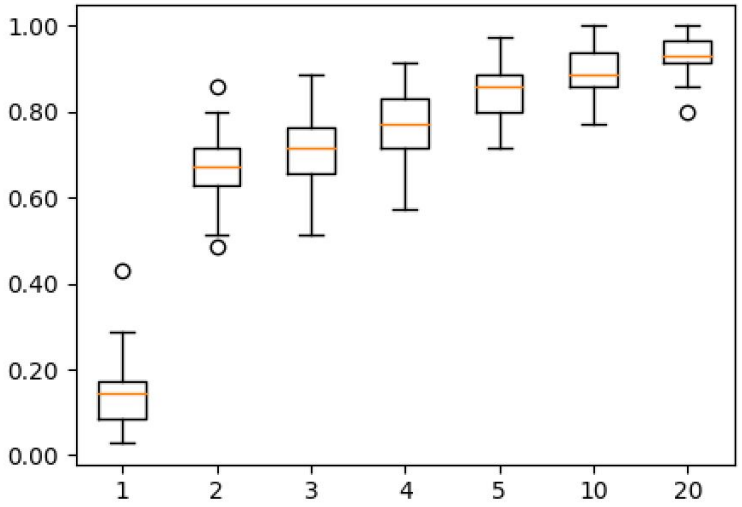
I.e. we needed to know what the target label was.
Need to ask for every class, and take highest if > 0.5

Accuracy



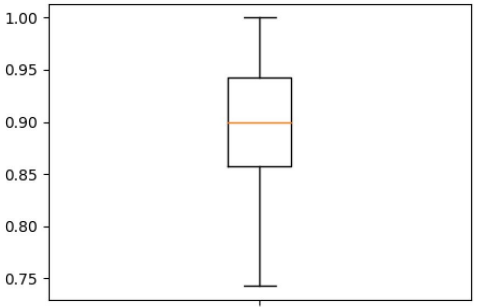
Reduced dimension embedding size

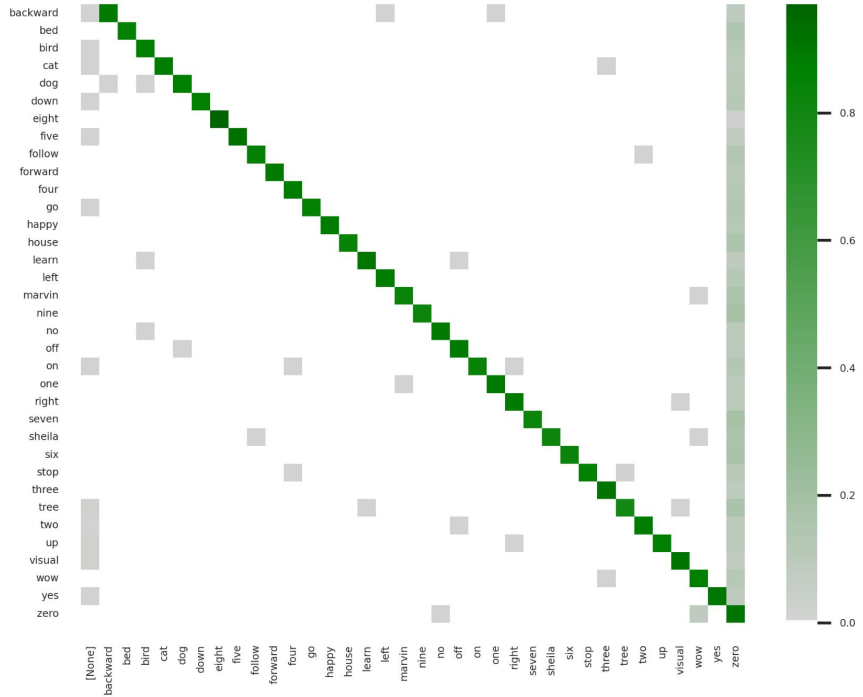
Accuracy



Number of examples

Confusion Matrix





Comparison with state-of-the-art

	Accuracy	
ANAN State of the Art	94%	de Andrade, D.C., Leo, S., Viana, M., Bernkopf, C.: A neural attention model for speech command recognition. ArXiv abs/1808.08929 (2018)
NUTS	89%	5 dimensions, **20** examples, AIKR limit 100
NUTS	64%	5 dimensions, **2** examples, AIKR limit 100

Open Questions

- Is there a magic ‘best’ matrix?
- Can it be found in a computationally efficient way?

Summary

It’s NUTS that **random** dimensionality reduction,

in combination with NARS,

works as well as it does.

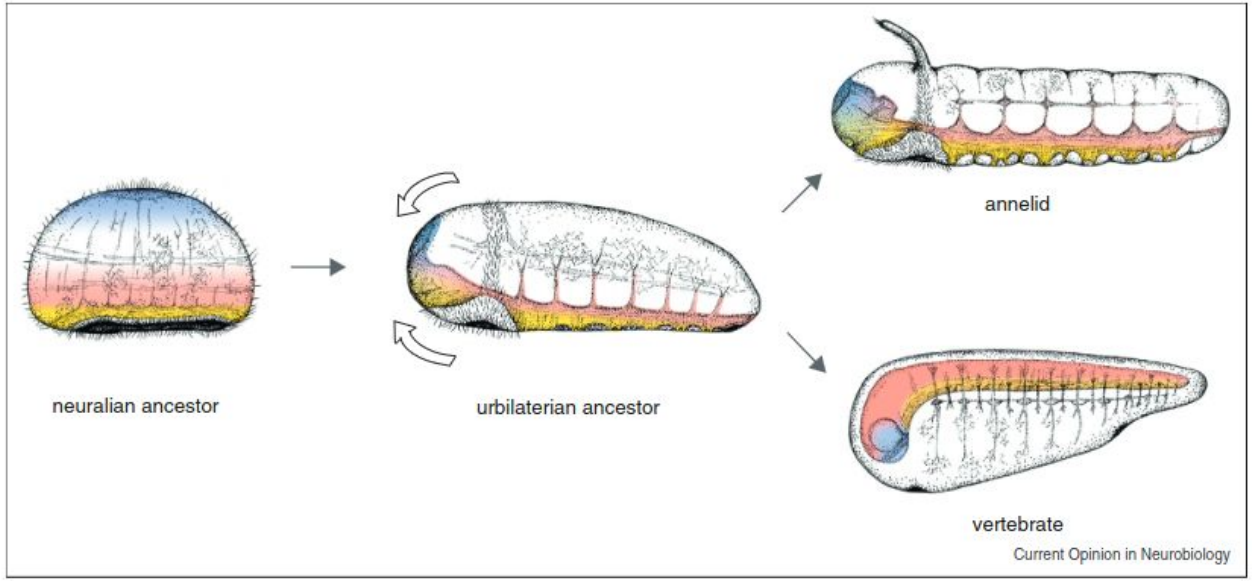
Appendix

Random Matrix

CS168: The Modern Algorithmic ToolboxLecture #4: Dimensionality Reduction

Peter from nPlan paper club suggested this: http://timroughgarden.org/s17/l/l4.pdf
Very good description on how multiplying by a random matrix can reduce dimensionality, yet keep (euclidean?) distance proportional within a desired epsilon https://dl.acm.org/doi/10.1145/1646353.1646379

Figure 3



The chimeric brain hypothesis. In the scheme, the apical nervous system is in blue, and the motor and the sensory-integrative components of the blastoporal nervous system are represented in yellow and red, respectively. In the last common ancestor of Cnidaria and Bilateria (neuralian ancestor) the two systems were separate, and concentrated respectively around the apical pole and the blastopore. With the transition to the urbilaterian (the ancestor of all Bilateria) the apical nervous system and anterior side of the blastoporal nervous system merged to form the forebrain, as found in extant annelids and vertebrates.

Source: Maria Antonietta Tosches, Detlev Arendt, The bilaterian forebrain: an evolutionary chimaera, Current Opinion in Neurobiology