R3.A.09

Real-Time Identification of Simple and Extended Musical Chords using Artificial Neural Networks

Coronel, Lesli Natasha A.
Navarro, Joachim Alfonso A.

Musical Chords

G
C

2 or more notes

E
C#
A

Played together

C
A
F#
D

Follow “rules of harmony”

(Leino, Brattico, Tervaniemi, & Vurst, 2007)

Musical Chords

BACKGROUND

Each has a name

C5
G
C

Amaj
E
C#
A

D7
C
A
F#
D

Musical Chords

Each has a root note

C5	Amaj	D7
		C
	E	A
G	C#	F#
C	A	D

Musical Chords

Each has a type

C5
G
C

Amaj
E
C#
A

D7
C
A
F#
D

Musical Chords

Simple vs
Extended

Chord types

Am
E
C
A

Simple

More common chord type

Musical Chords

Simple vs Extended

Chord types

AmM7

G#

E

C

A

Extension

Extended

Less common chord type

The determination of the name of the chord from the notes that constitute it

Definition of chord identification

Chord Identification

PROBLEM

Majority of general music learning public **can't do this by themselves** due to **lack of skill** or training

Situation with chord identification

Neural networks

DEFINITION

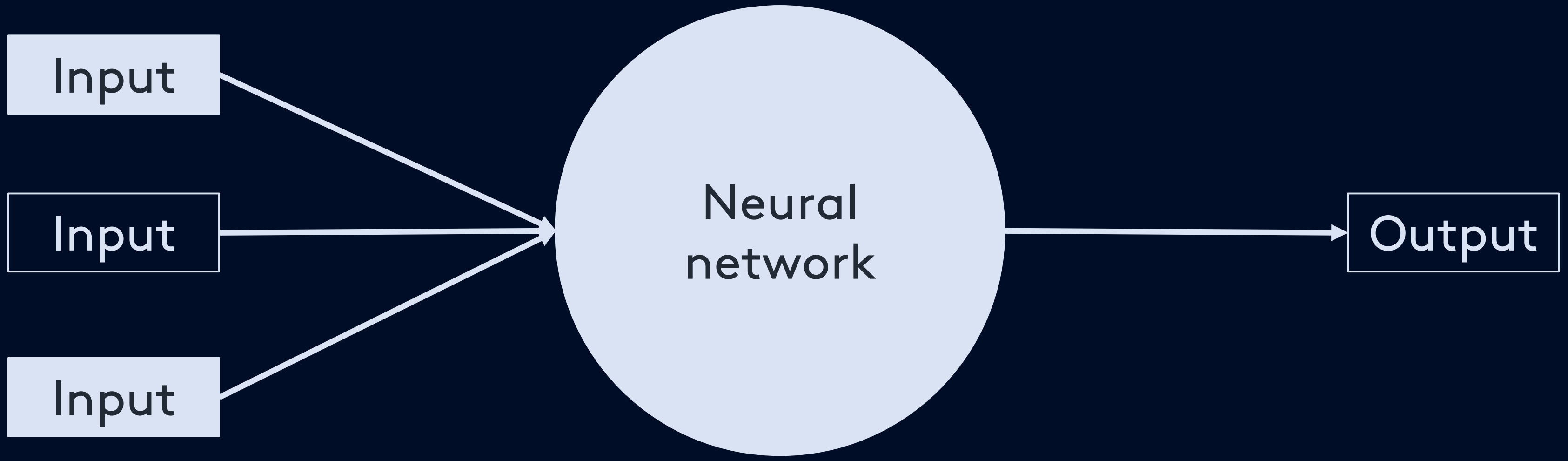
Computational model of neurons in a brain

Neural networks

DEFINITION

Many neurons = neural network

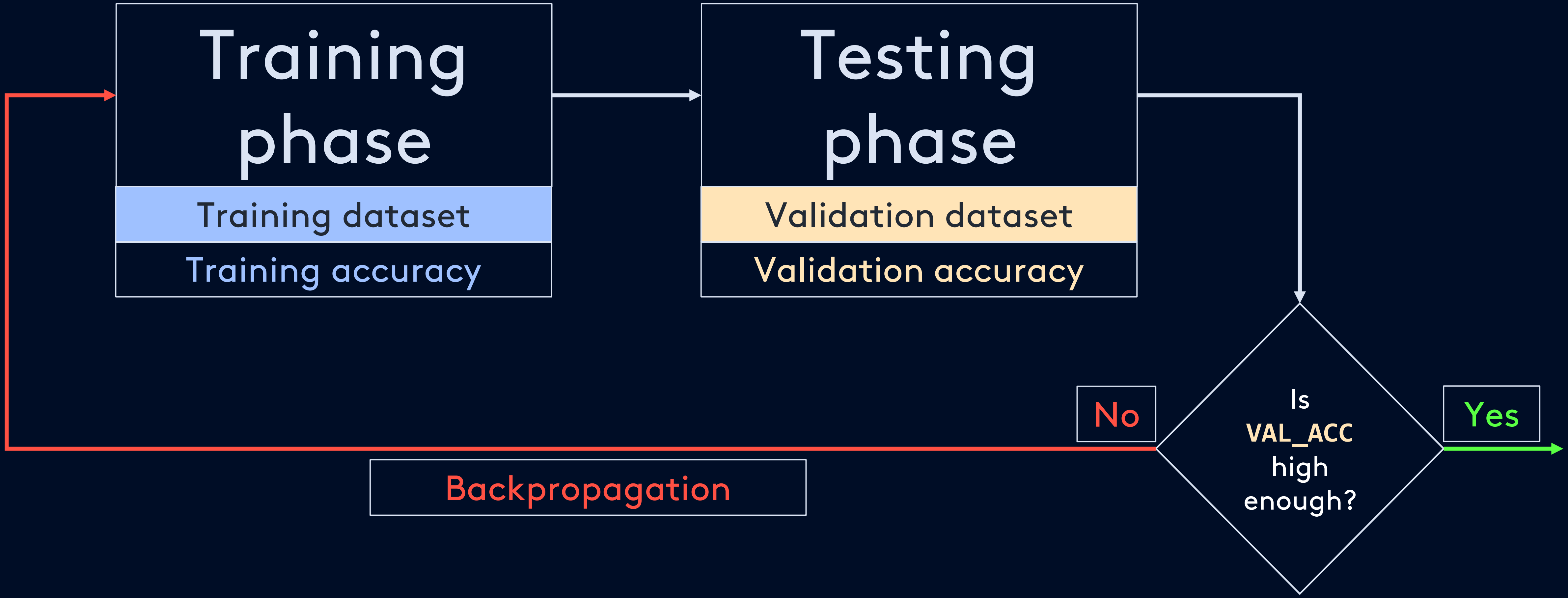
Artificial Neural Networks (ANNs)



ANN learns by repetitive training

Colina, Perez, & Paraan, 2017

ANN training & testing



Why neural networks?

Previous studies with neural network implementations have **not included extended chords in their research**

Osmalskyj, Embrechts, Piérard, & Van Droogenbroeck, 2012
Perera & Kodithuwakku, 2005
Zhou & Lerch, 2015

Problem statement

Using neural networks to identify both common and extended chords is unexplored

Osmalskyj, Embrechts, Piérard, & Van Droogenbroeck, 2012
Perera & Kodithuwakku, 2005
Zhou & Lerch, 2015

Develop a neural network that quickly identifies simple and extended musical chords

Input is a group of 3 or more MIDI note signals played in real-time

Minor

Input chords have **one root note** and are **not inverted**

Identification must be quick enough to be used in live performance (<40ms)

Greeff, 2016

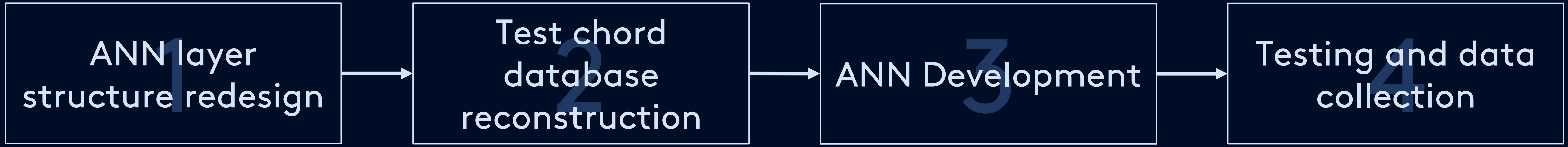
Implemented in programming languages with **neural network**, **real-time MIDI**, and **GPU processing** libraries

thestk, 2017; Bretschneider, 2017

Neural network must be run on a GPU for efficient processing

Nickolls, Buck, Garland, & Skadron, 2008

Level 0


1
ANN layer structure redesign
2
Test chord database reconstruction
3
ANN Development
4
Testing and data collection

Performance

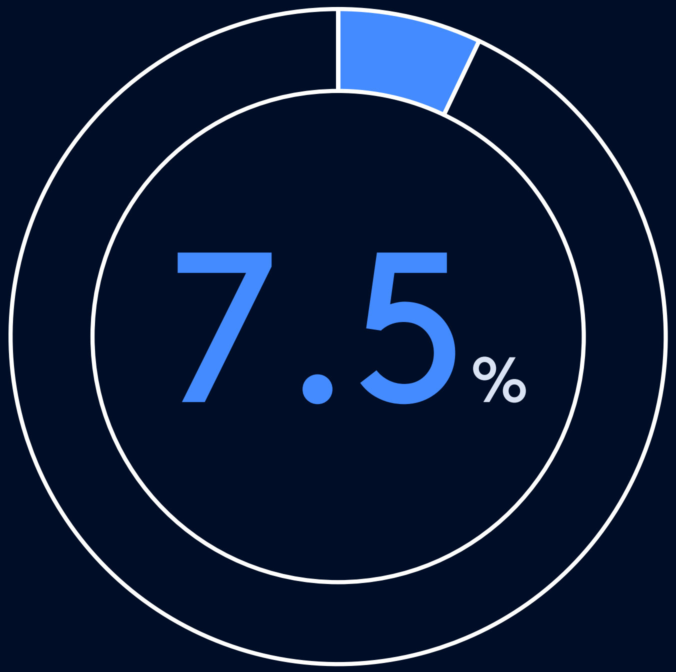
Peak validation accuracy after 30K epochs

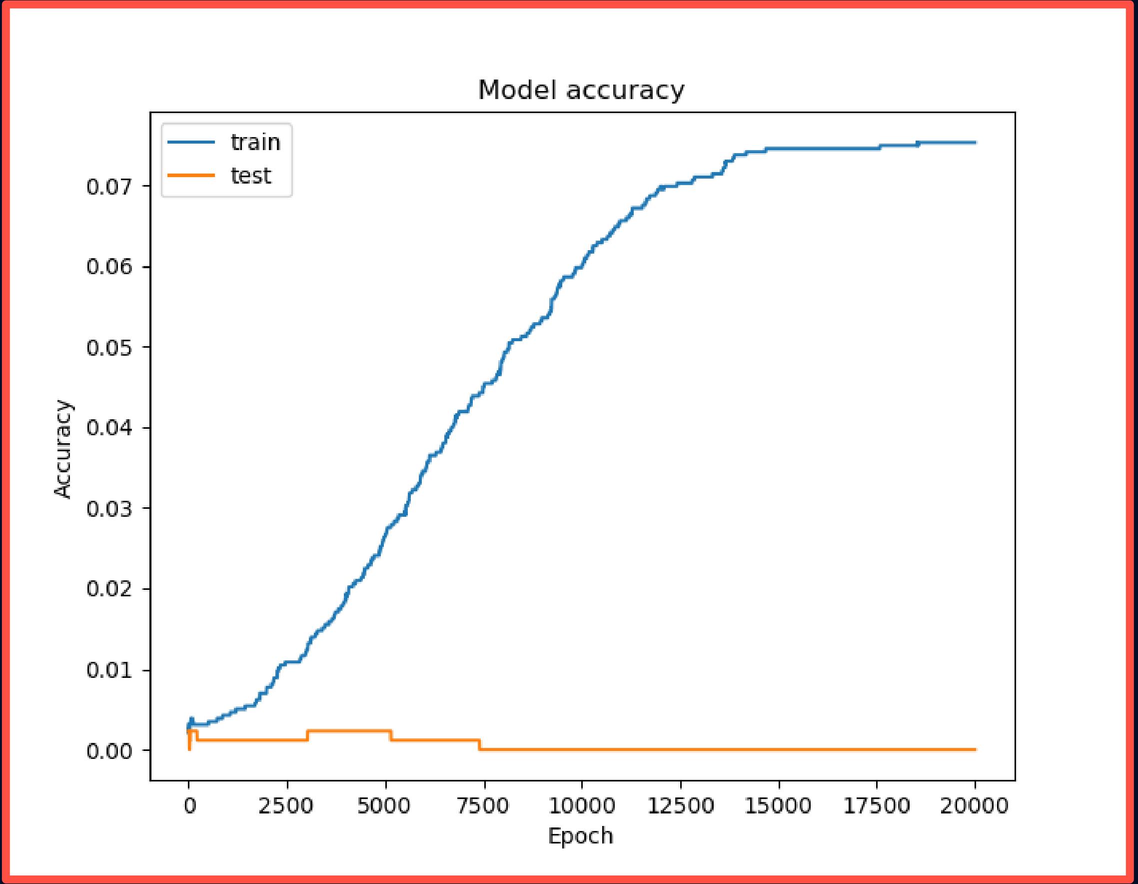
Peak training accuracy after 30K epochs

Coronel
Navarro

Performance

Peak training accuracy
after 30K epochs



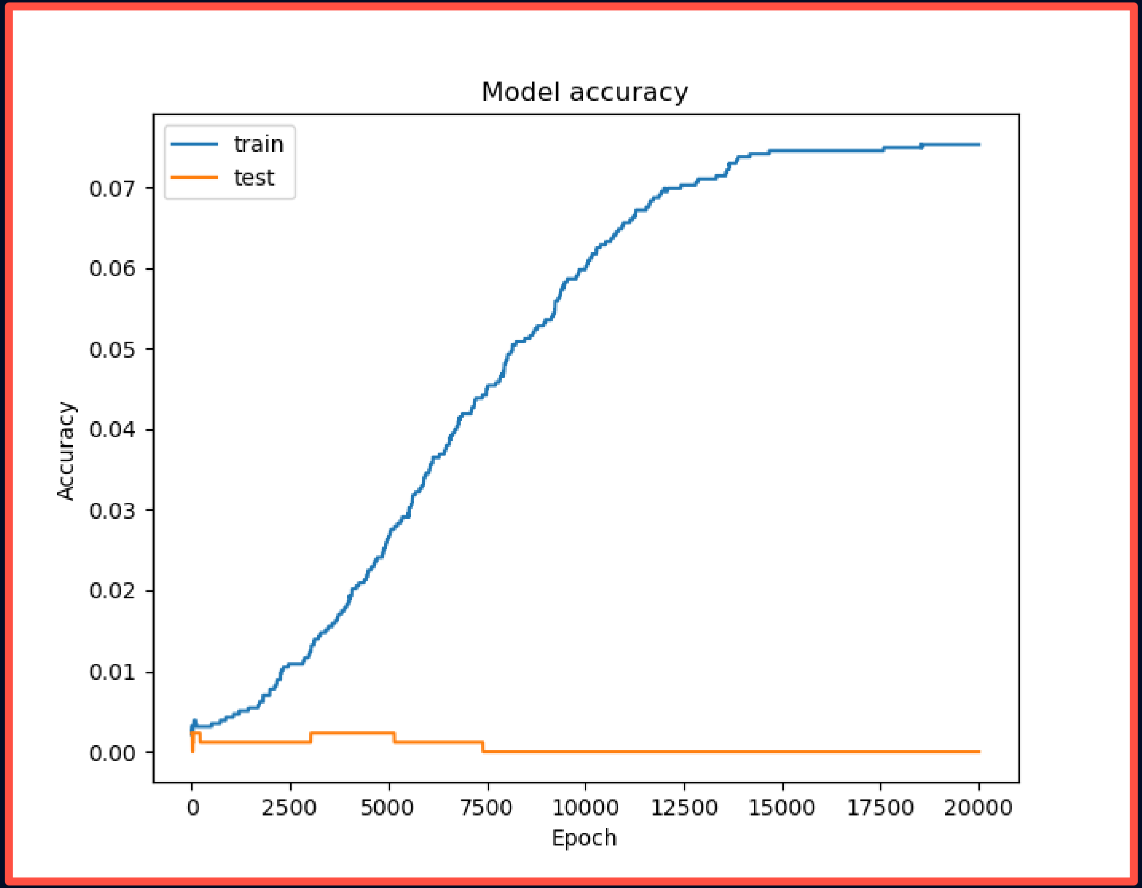


Plateaus on learning **training dataset**

Trend

Performance

Peak validation accuracy
after 2800 epochs



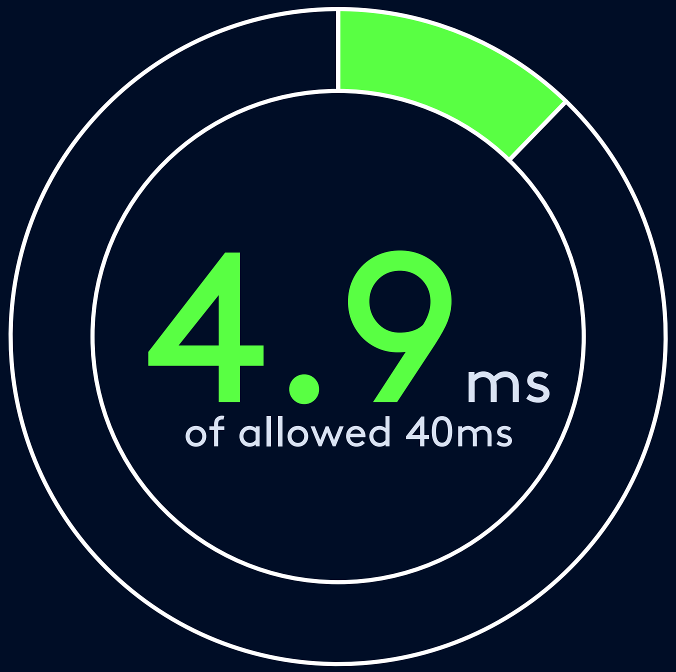
Over-fitting on training dataset

Reason

Coronel
Navarro

Performance

Mean total response time,
30 samples



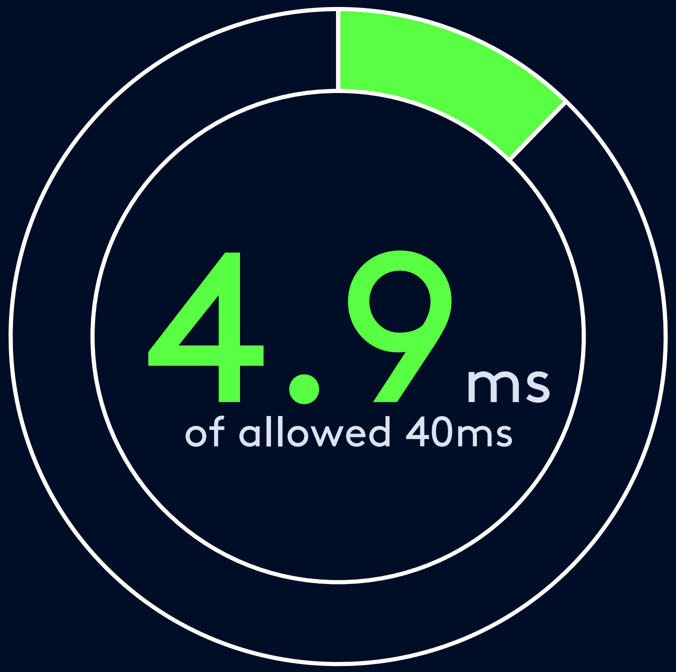
Null hypothesis	Alternative hypothesis
$r \geq 40ms$	$r < 40ms$

T-test for one mean

Sample size = 30; Significance = 5%

Performance

Mean total response time,
30 samples



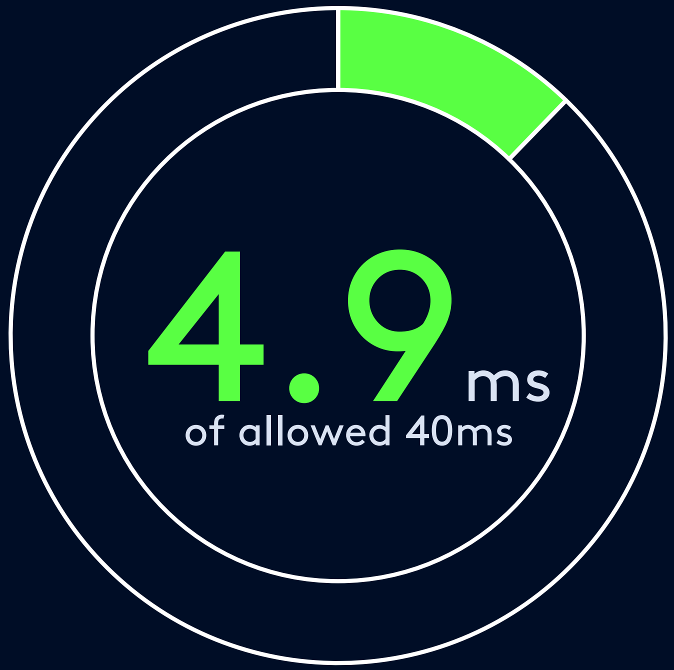
Null hypothesis	Alternative hypothesis
$t \geq -1.699$	$t < -1.699$

T-test for one mean
Sample size = 30; Significance = 5%

Performance

RESULTS

Mean total response time,
30 samples



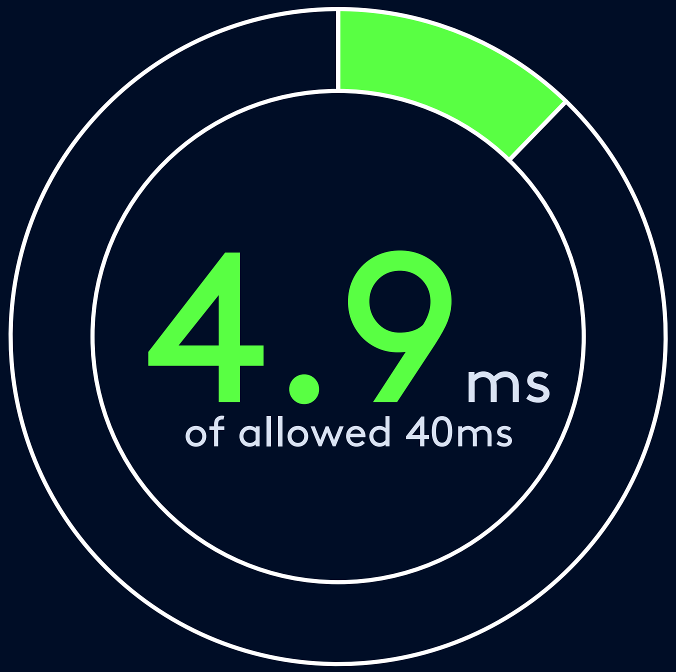
Null hypothesis	Alternative hypothesis
$t \geq -1.699$	$t < -1.699$
$t_{4.9ms} = -3.17$	

T-test for one mean

Sample size = 30; Significance = 5%

Performance

Mean total response time,
30 samples



Null hypothesis	Alternative hypothesis
$t \geq -1.699$	$t < -1.699$

REJECT

NN is faster than standard!

T-test for one mean
Sample size = 30; Significance = 5%

Conclusion

Our chords are too complex for NN...

...but NNs are fast enough

Recommendations

1	2	3
Other machine learning algorithms	Simplified set of chords	Use audio rather than MIDI as input

T	H	E	Rev. C1	E	N	D

Thank you!

References

Baharloo, S., Service, S., Risch, N., Gitschier, J., & Freimer, N. (2000). Familial aggregation of absolute pitch. *American Journal of Human Genetics, 67, 755-758.* doi:10.1086/303057.

Chord. (2004). In *The concise Oxford dictionary of music* (4th ed.) Oxford, UK: Oxford University Press.

Colina, N. C. A., Perez, C. E., & Paraan, F. N. C. (2017). Simple techniques for improving deep neural network outcomes on commodity hardware. *AIP Conference Proceedings*, *1871*, 040001. doi:10.1063/1.4996523.

Daniel, G. (2013). *Principles of artificial neural networks* (3rd ed.) Chicago, IL: World Scientific.

Fujishima, T. (1999). Realtime chord recognition of musical sound: A system using common Lisp music. Retrieved from http://www.music.mcgill.ca/~jason/mumt621/papers5/fujishima_1999.pdf .

Greeff, W. (2016). The influence of perception latency on the quality of musical performance during a simulated delay scenario. Retrieved from https://repository.up.ac.za/bitstream/handle/2263/58578/Greeff_Influence_2017.pdf?sequence=4.

Humphrey, E., Bello, J. P., & Cho, T. (n.d.). Chord Recognition. Retrieved from http://steinhardt.nyu.edu/marl/research/chord_recognition.

Kidd, P. (2017). pyrtmidi: Real-time MIDI I/O for Python [GitHub repository]. Retrieved August 23, 2017, from https://github.com/patrickkidd/pyrtmidi.

Nickolls, J., Buck, I., Garland, M., & Skadron, K. (2008). Scalable parallel programming with CUDA. *ACM Queue*, *6*(2), 40-53.

Osmalskyj, J., Embrechts, J-J., Piérard, S., & Van Droogenbroeck, M. (2012, May 9). Neural networks for musical chords recognition. Retrieved at http://jim.afim-asso.org/jim12/pdf/jim2012_08_p_osmalskyj.pdf.

Perera, N., & Kodithuwakku, S. R. (2005, December 15). Music chord recognition using artificial neural networks. *1st Proceedings of the International Conference on Information and Automation*, 304-308.

Root. (2004). In *The concise Oxford dictionary of music* (4th ed.) Oxford, UK: Oxford University Press.

Stark, A. M., & Plumbley, M. D. (2009). Real-time chord recognition for live performance [PDF file]. Retrieved at https://www.eecs.qmul.ac.uk/~markp/2009/StarkPlumbley09-icmc.pdf.

Zatorre, R. J., Perry, D. W., Beckett, C. A., Westbury, C. F., & Evans, A. C. (1998). Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proceedings of the National Academy of Sciences*, *95*, 3172-3177. Retrieved at http://www.pnas.org/content/95/6/3172.full.

Zhou, X., & Lerch, A. (2015). Chord detection using deep learning. *16th International Society for Music Information Retrieval Conference*, 52-58. Retrieved at http://ismir2015.uma.es/articles/96_Paper.pdf.