

# Exploring hyperparameters

RNN for generating Estonian text

Robert Roosalu

# Introduction

Replicate examples in Andrej Karpathy's blog post for Estonian

Keras example works great...

Explore hyperparameters!

# Dataset

160 books from Estonian literature (1990-...) → 35 MB

Time constraint

Results in 12 hours → 1 MB

Rein Põder "Hiliskevad" + 0.24 Ene Mihkleson "Ahasveeruse Uni"

# Model



2-layer RNN(LSTM)

Softmax activation

Categorical Crossentropy loss

Mini-batch Gradient Descent

Dropout

RMSPProp

# Hyperparameters

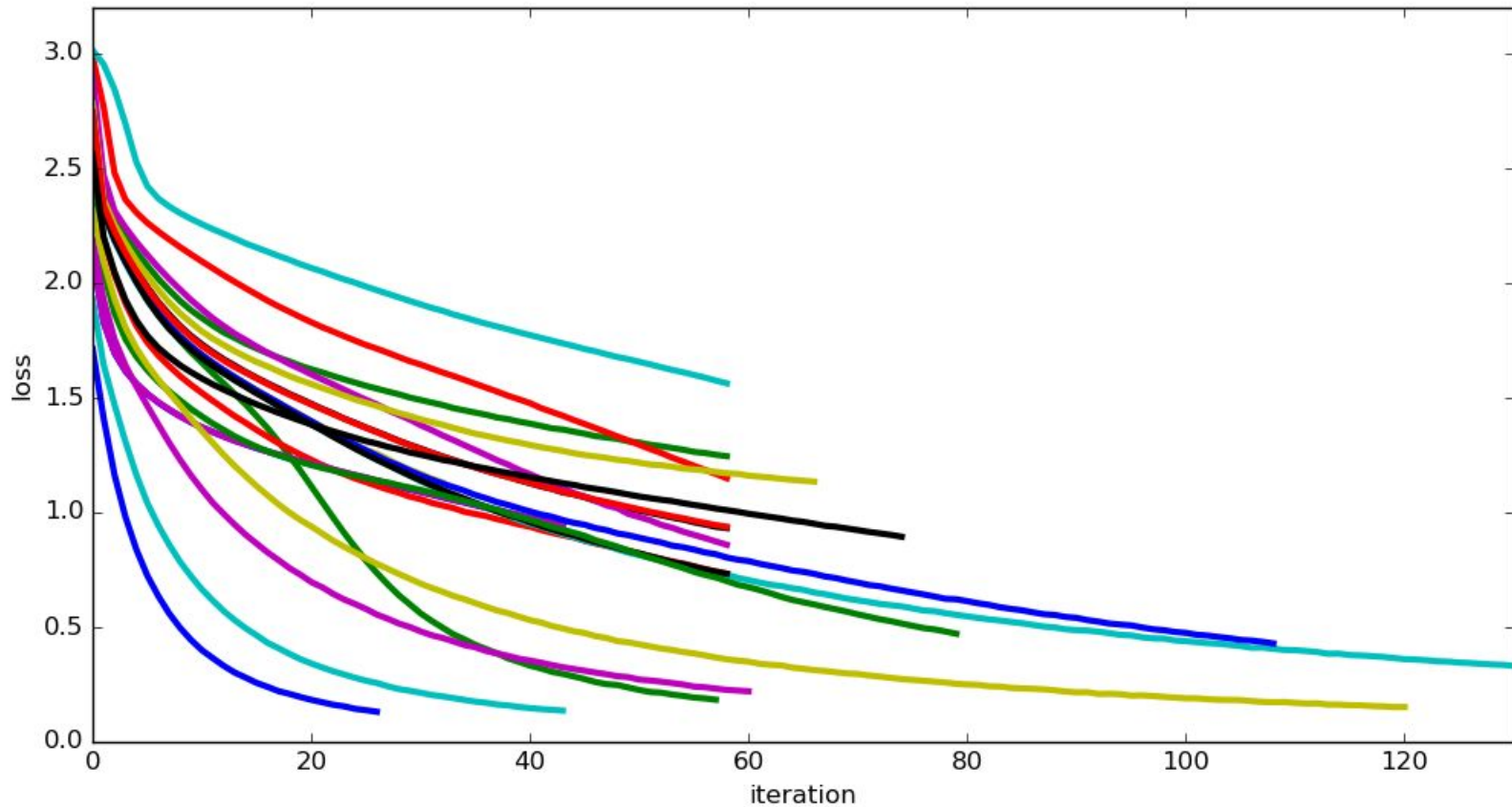


dataset: corpus, window\_size

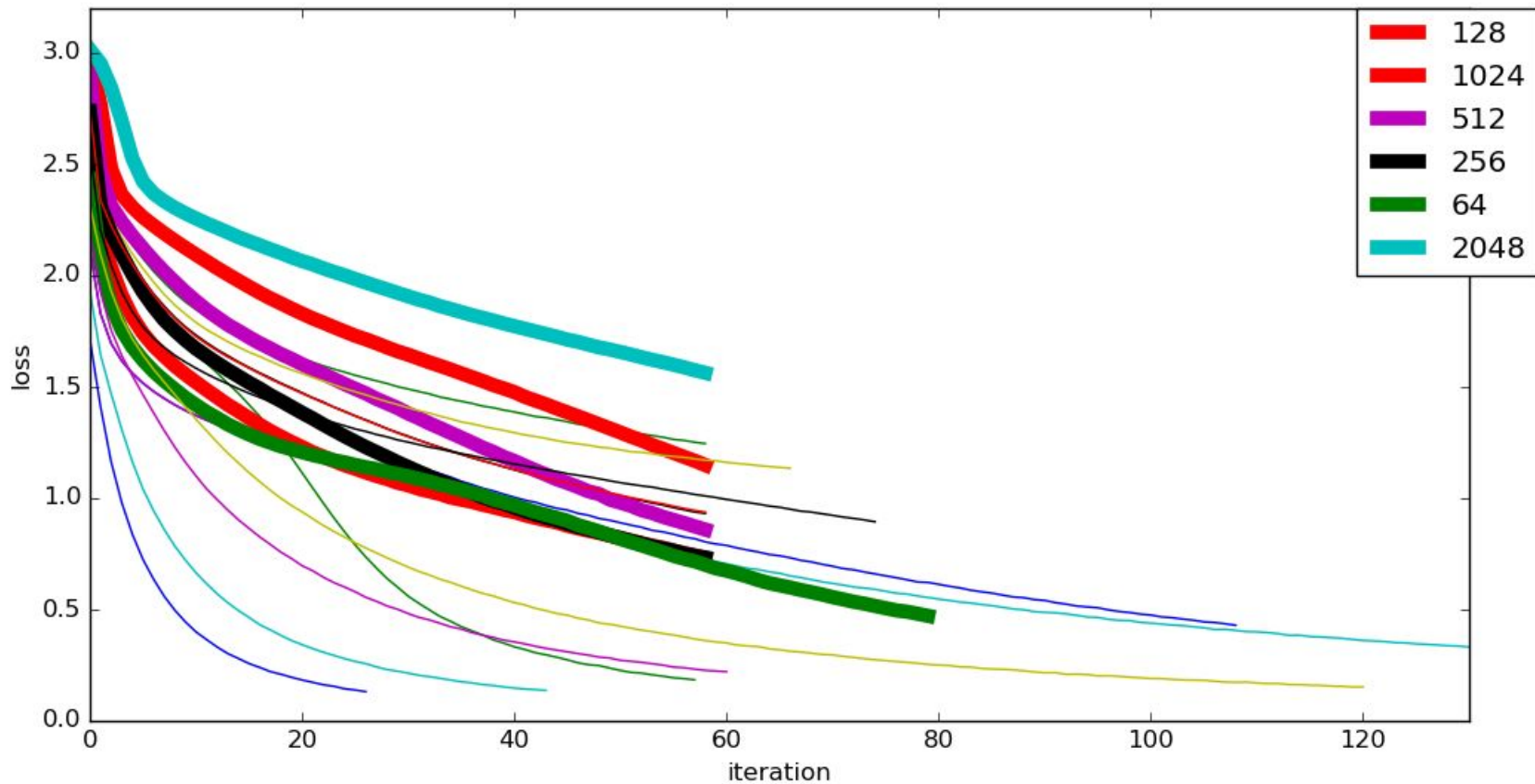
net: dropout, neuron\_count

optimisation: epoch\_number, batch\_size

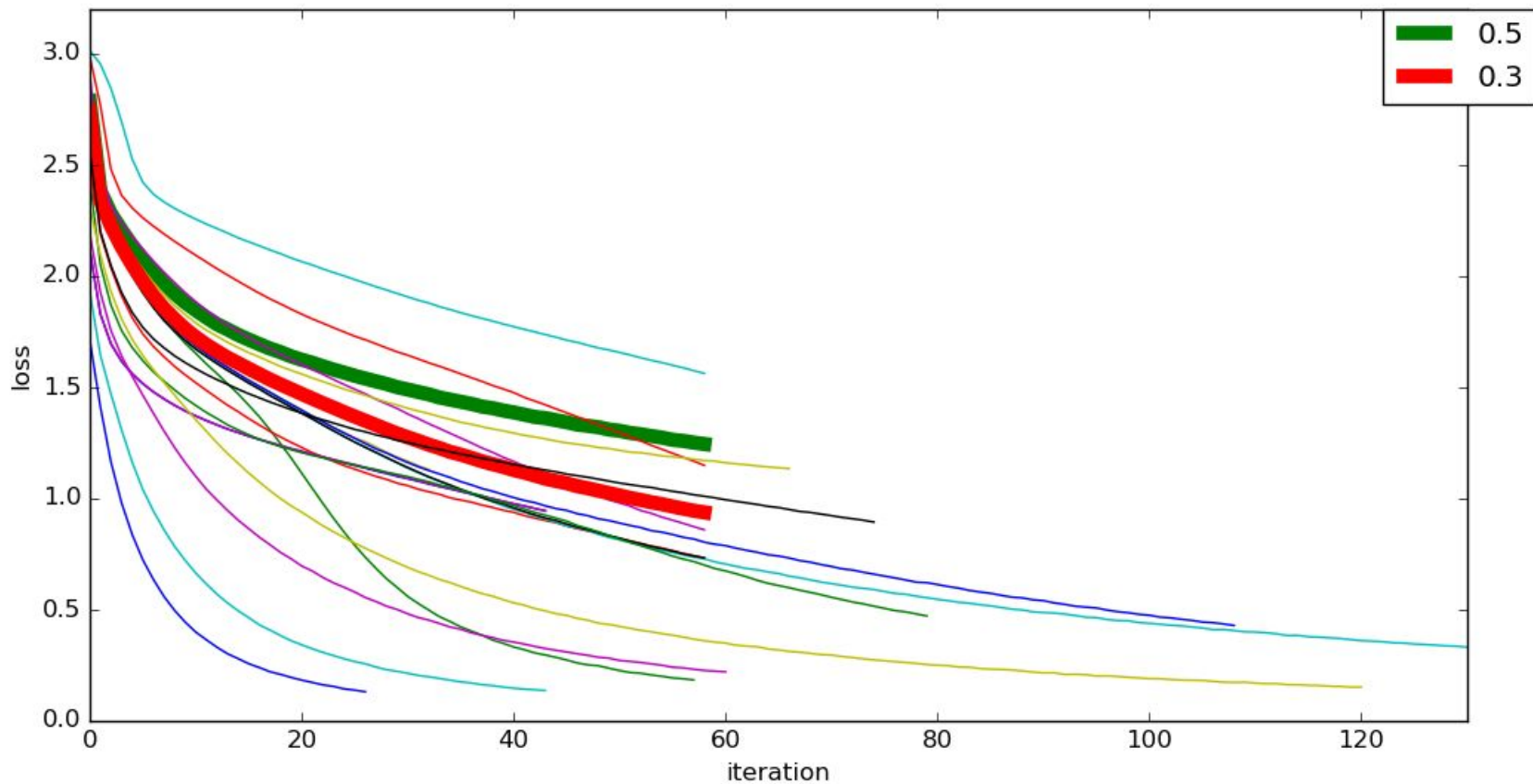
# Cost functions



# Batch size

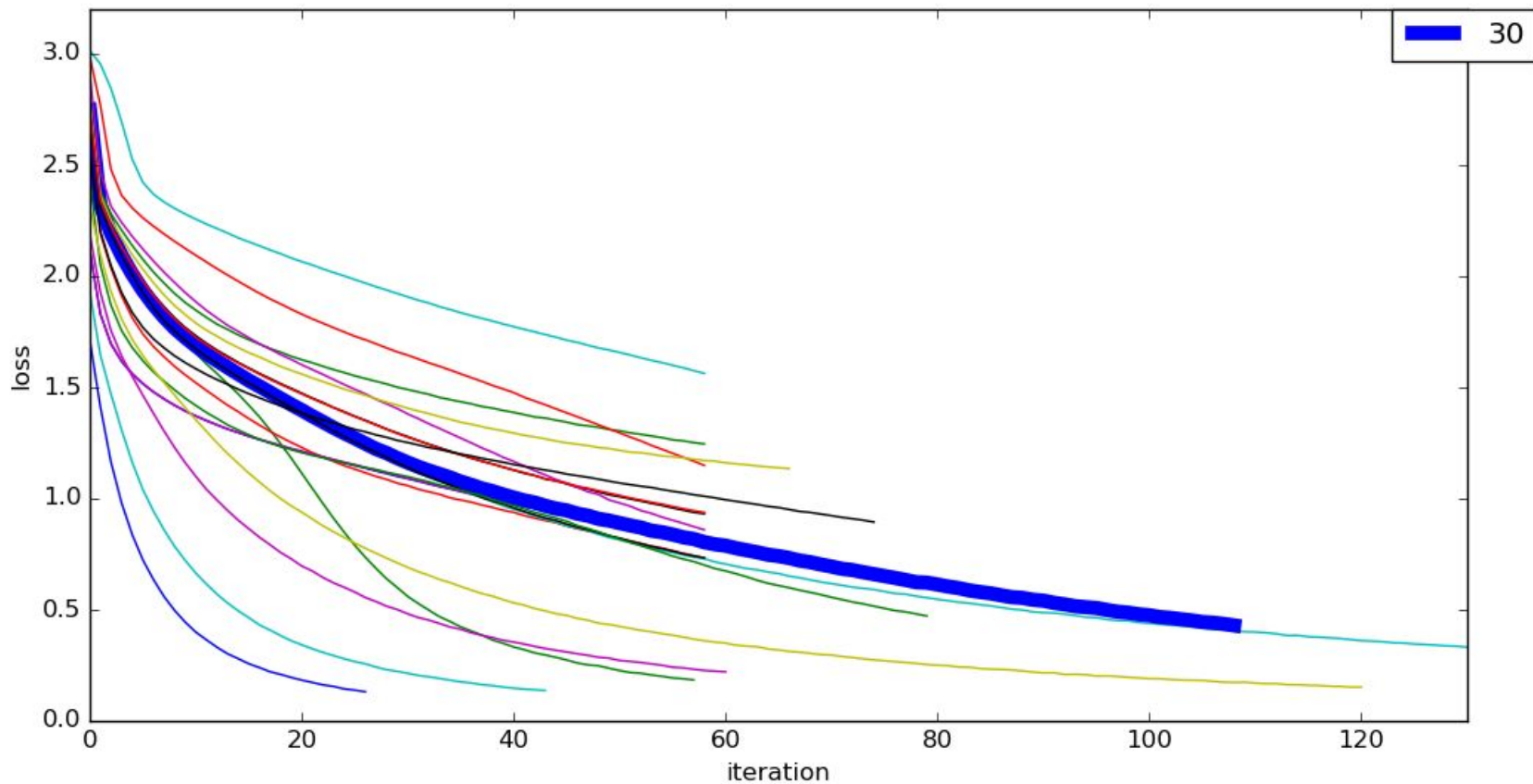


# Dropout

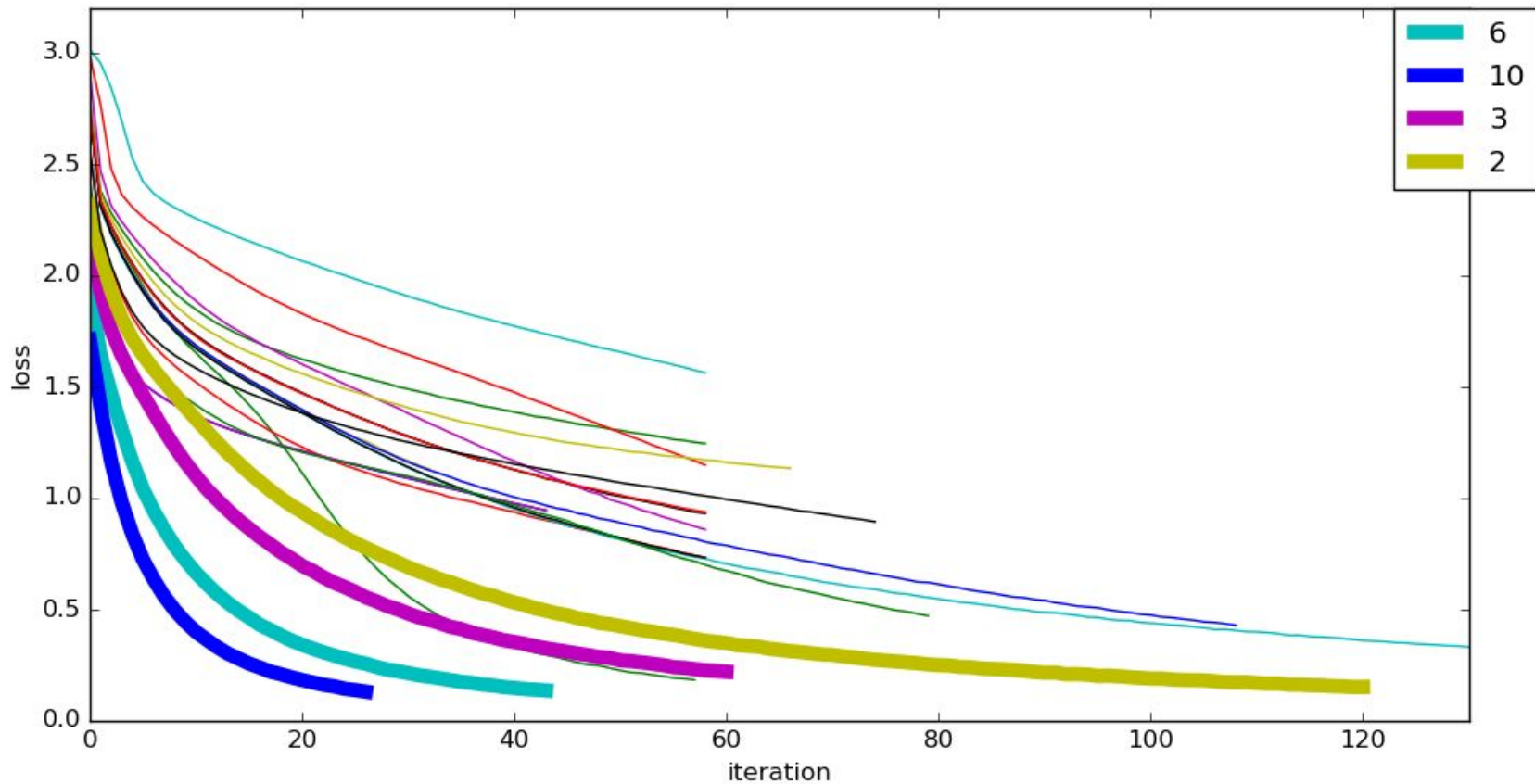




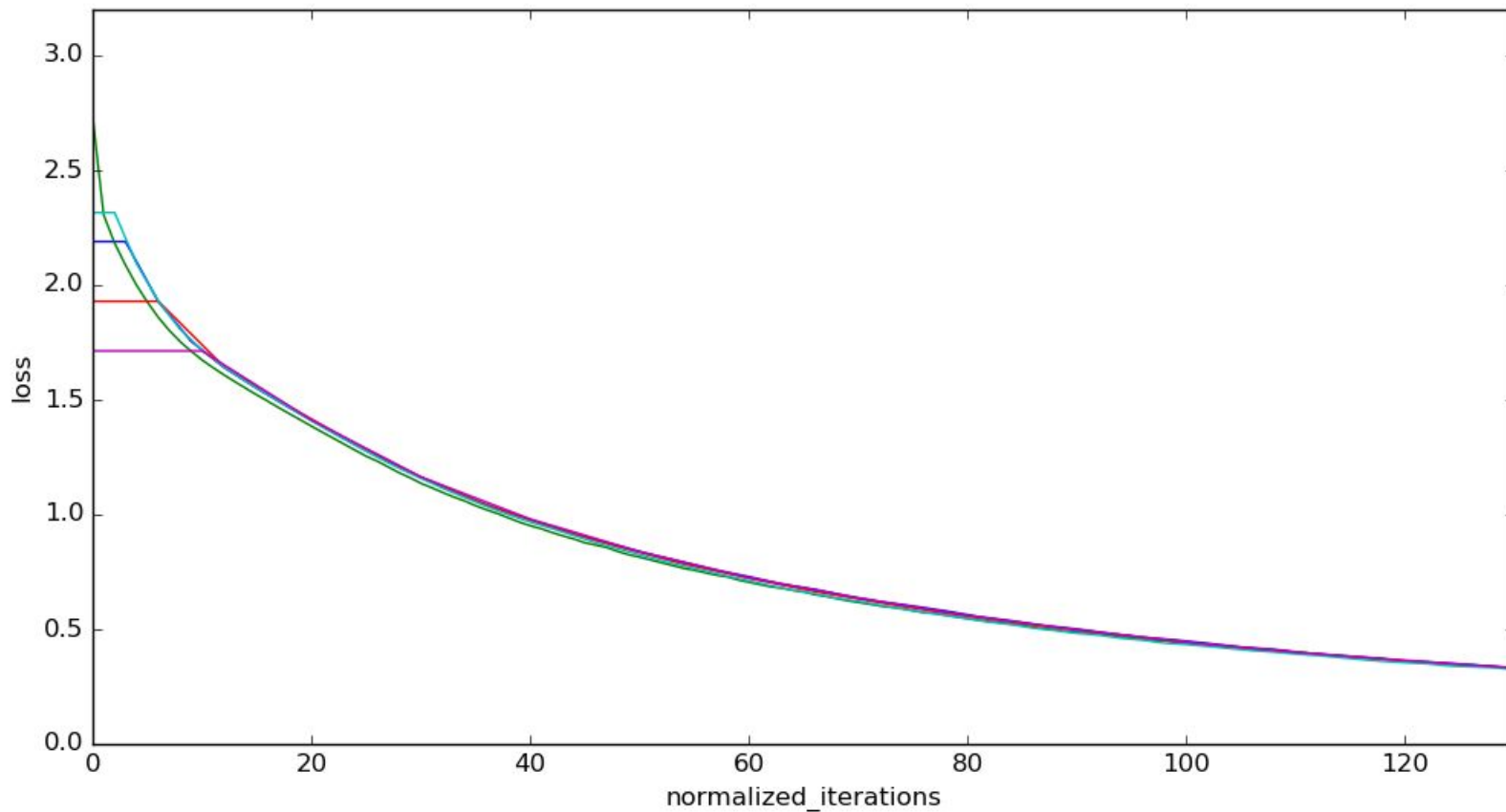
# Window size



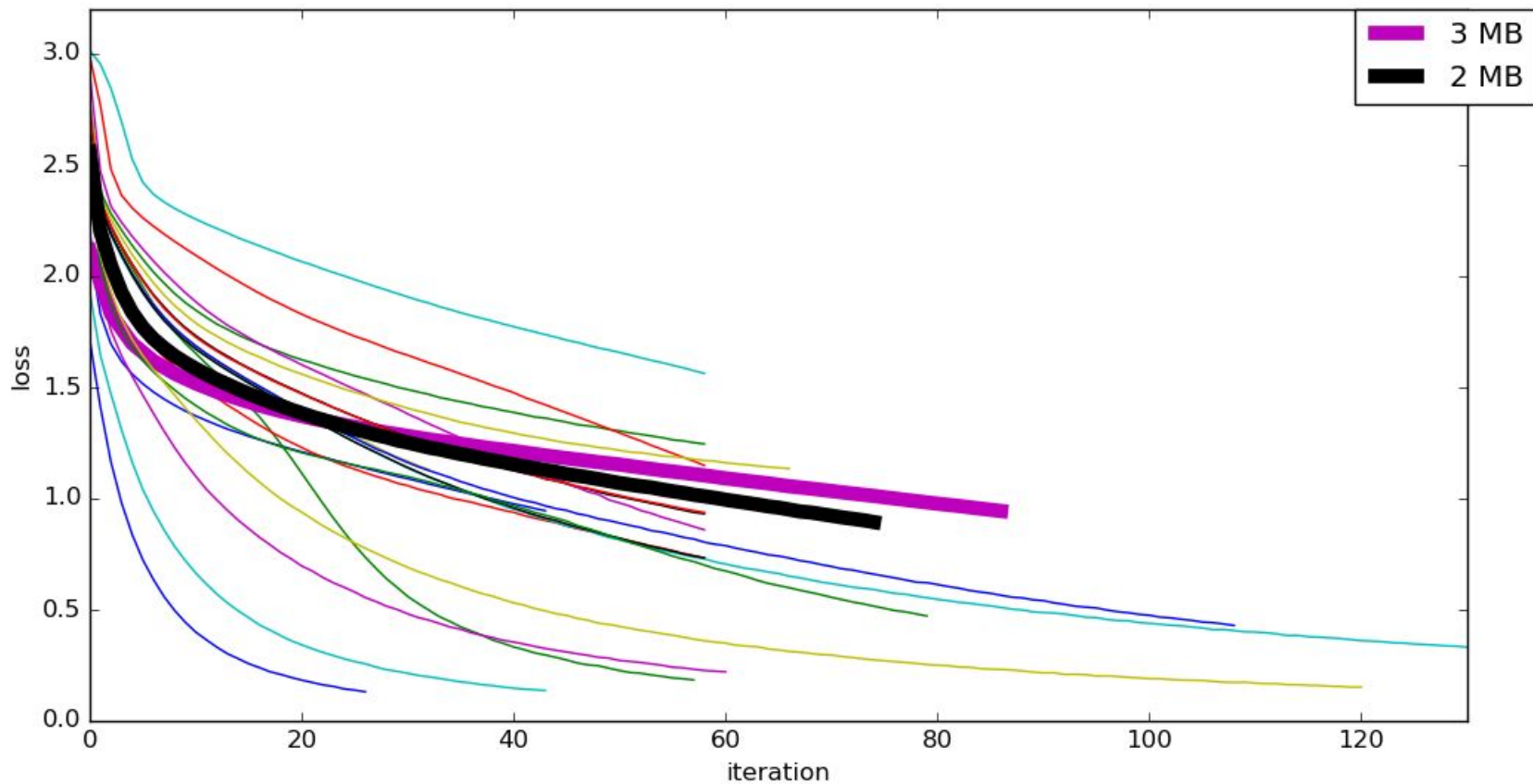
# Epoch



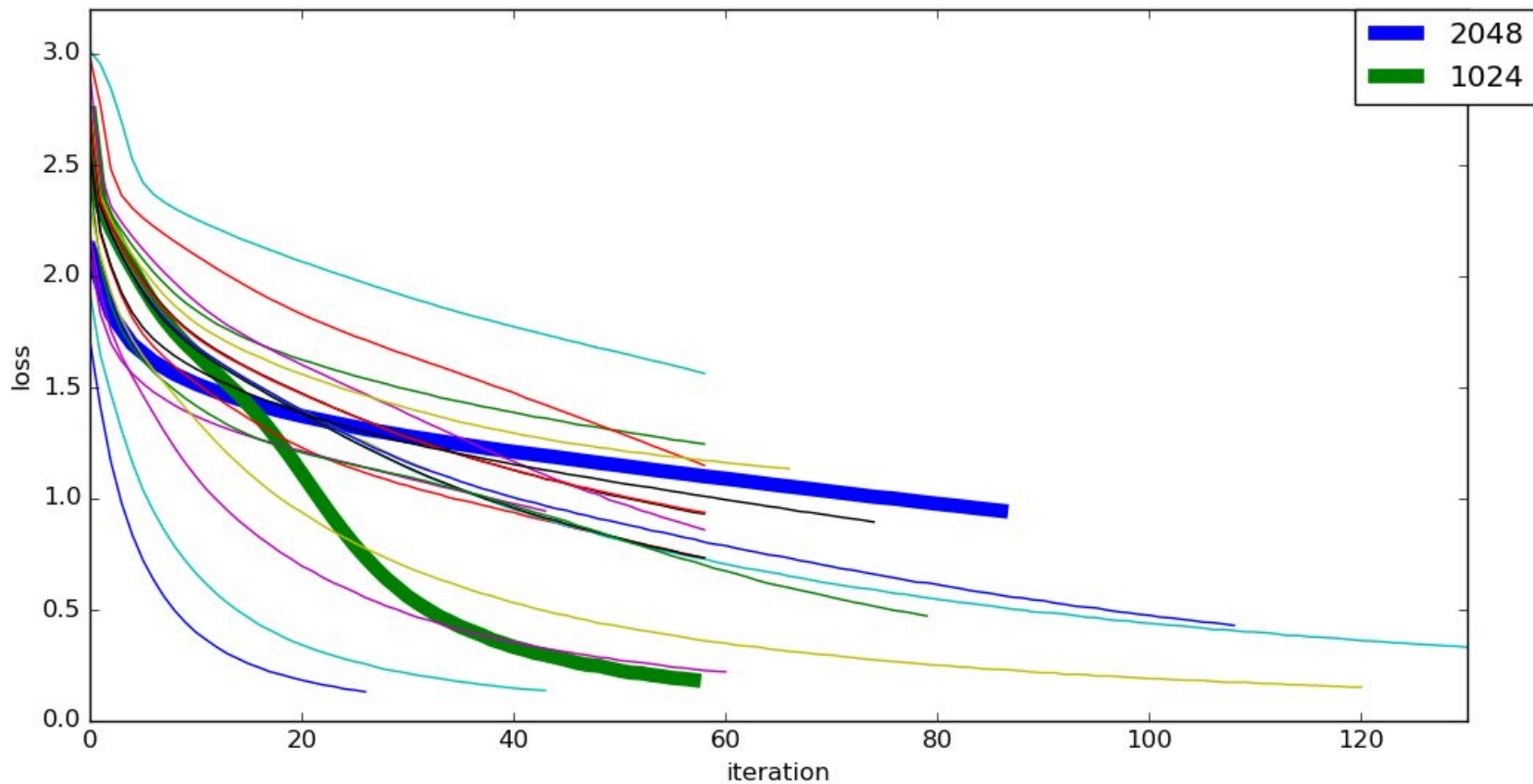
# Epoch == iteration



# Corpus size



# Neurons

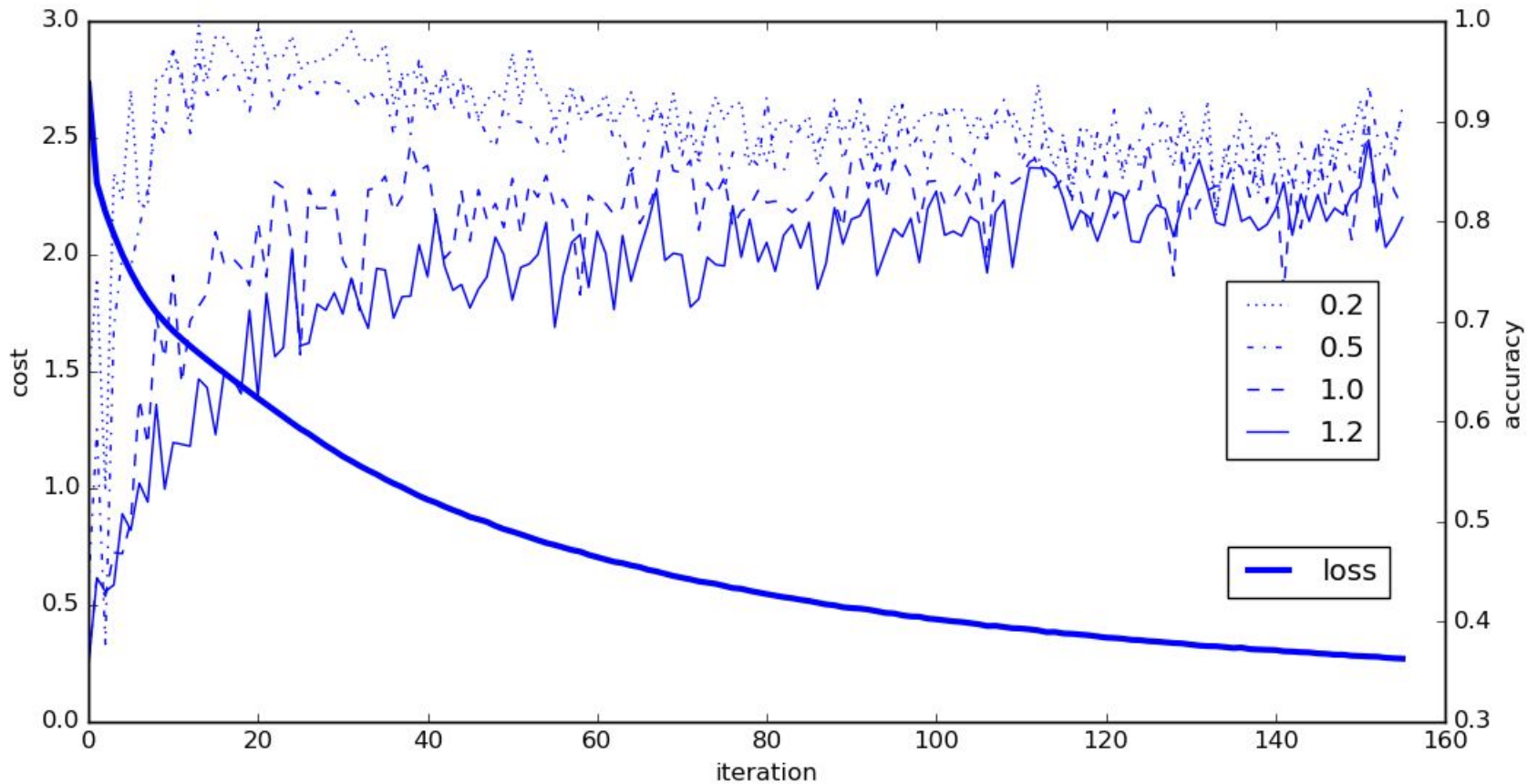


# Measuring accuracy



$$\textit{accuracy}_{temp} = \frac{\textit{correct words}}{\textit{generated words}}$$

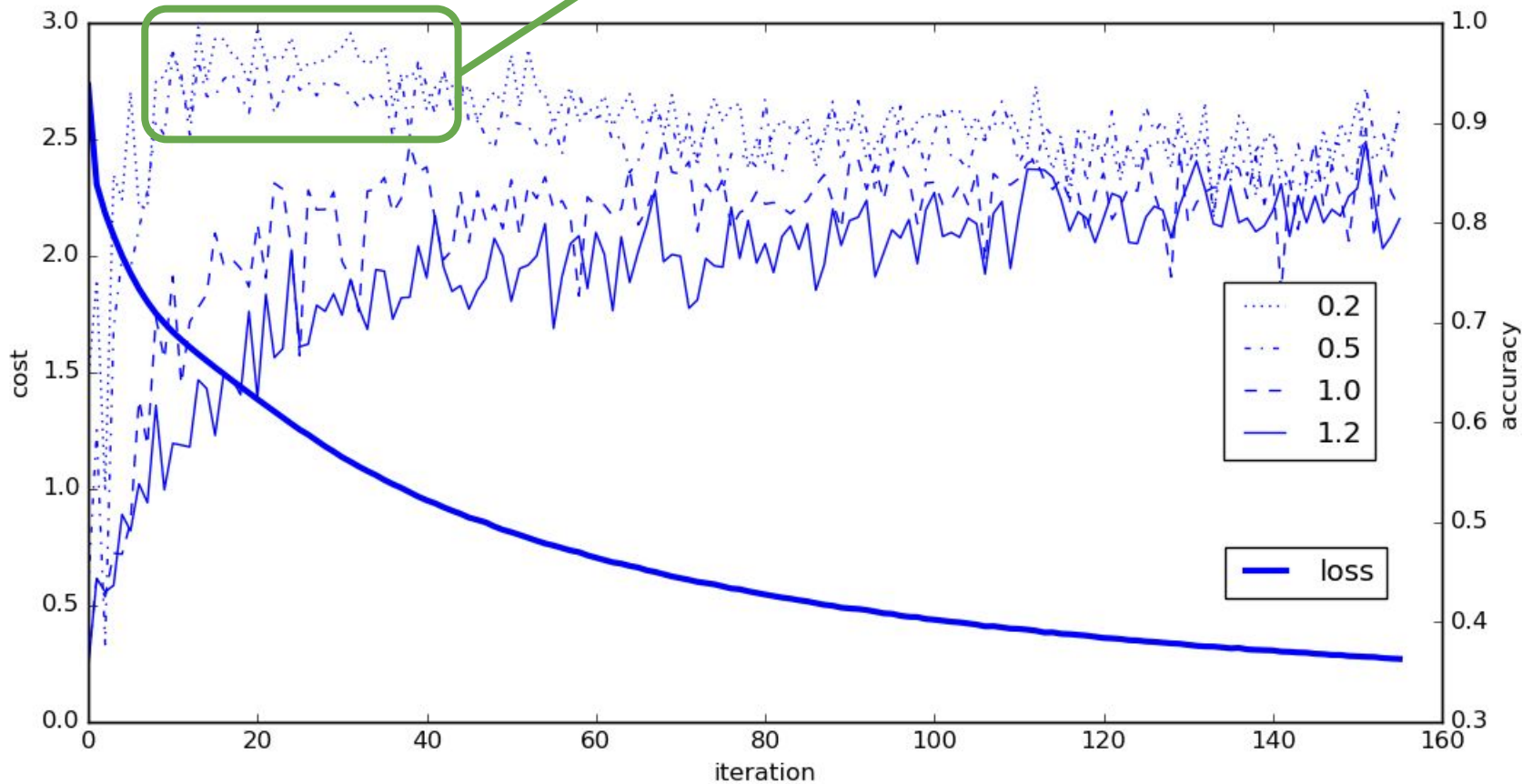
# Measuring accuracy





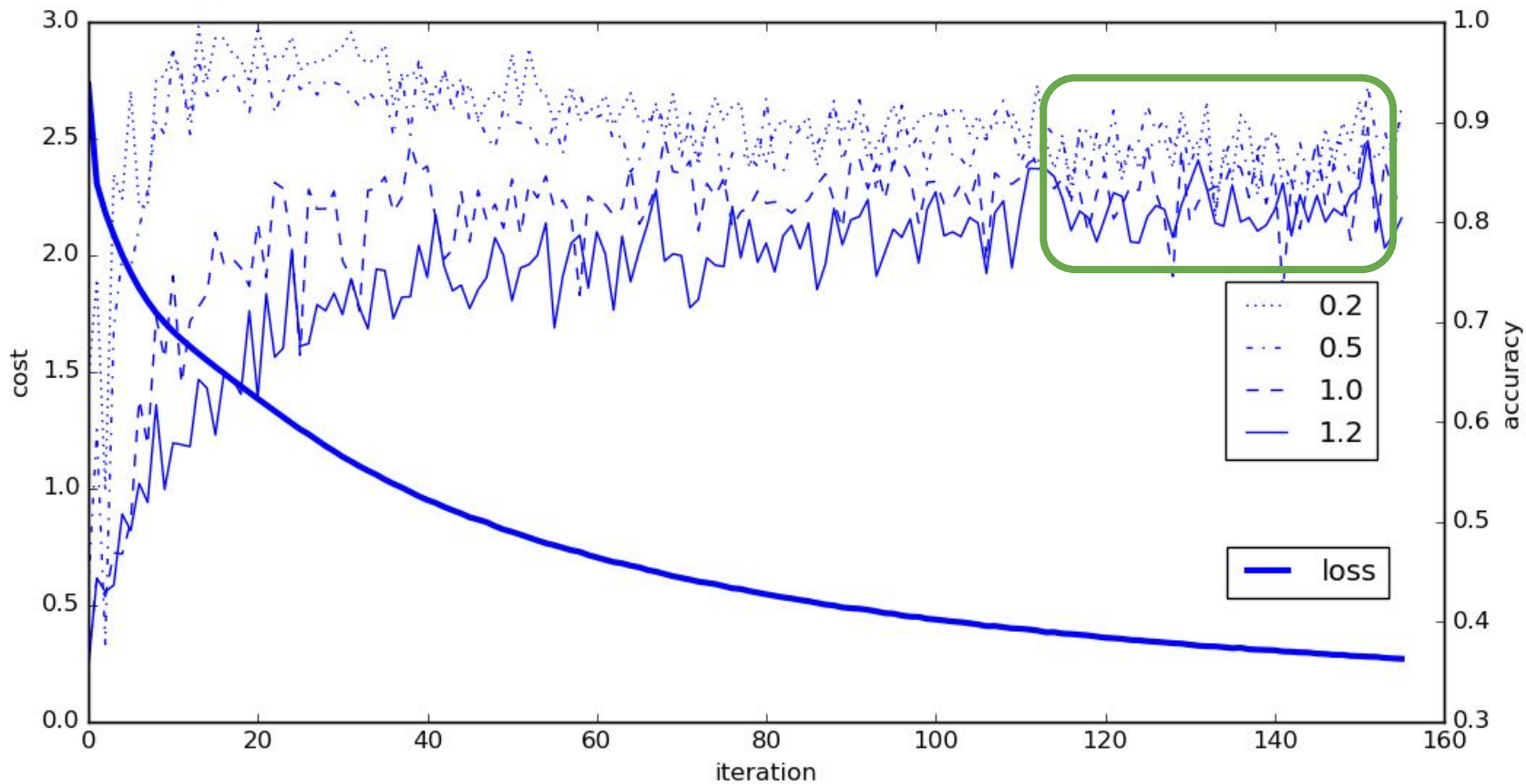
# Measuring accuracy

ikka seda üksijäri ma oli mu ei ta ning  
telle vae oli neid kui ta minu

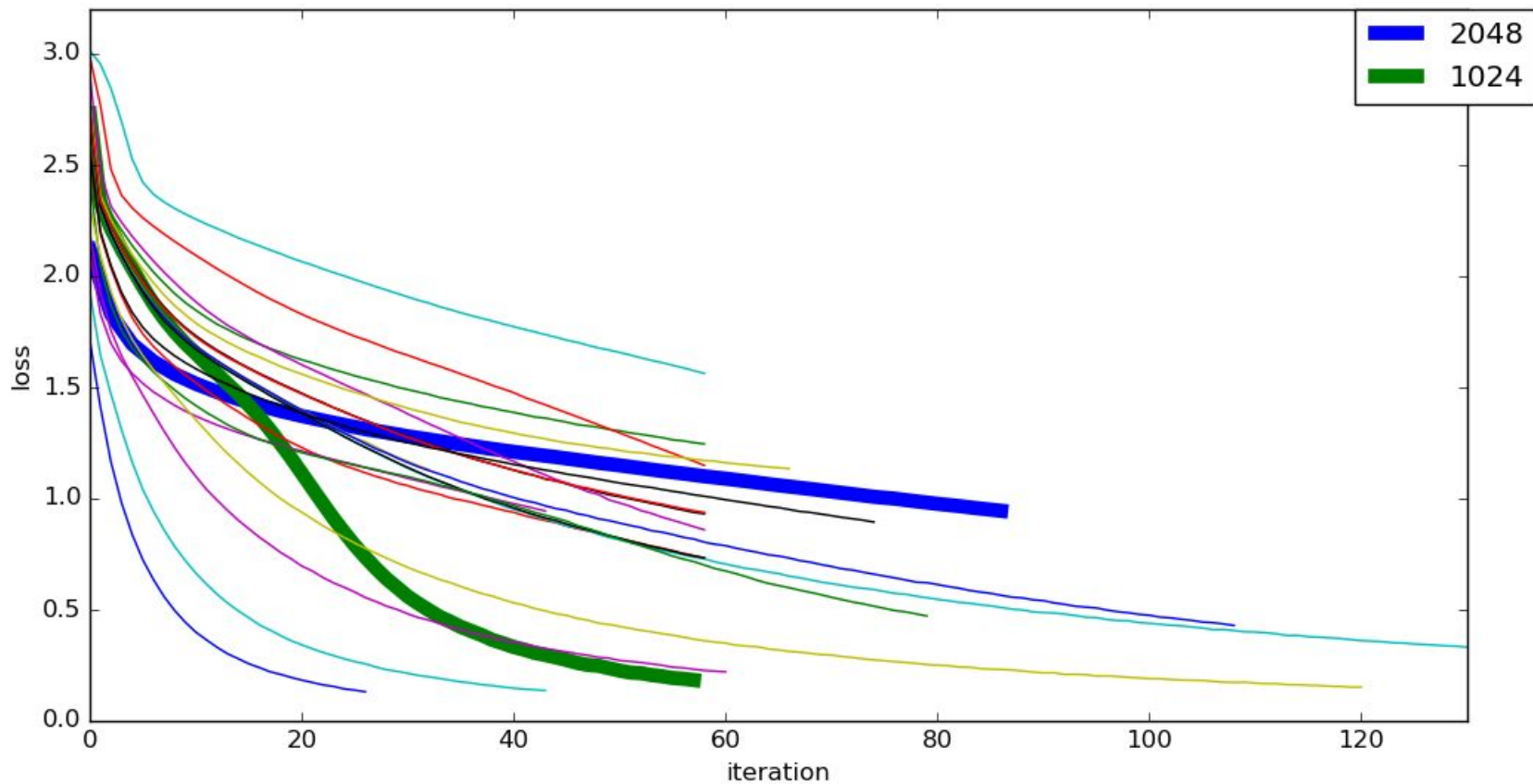




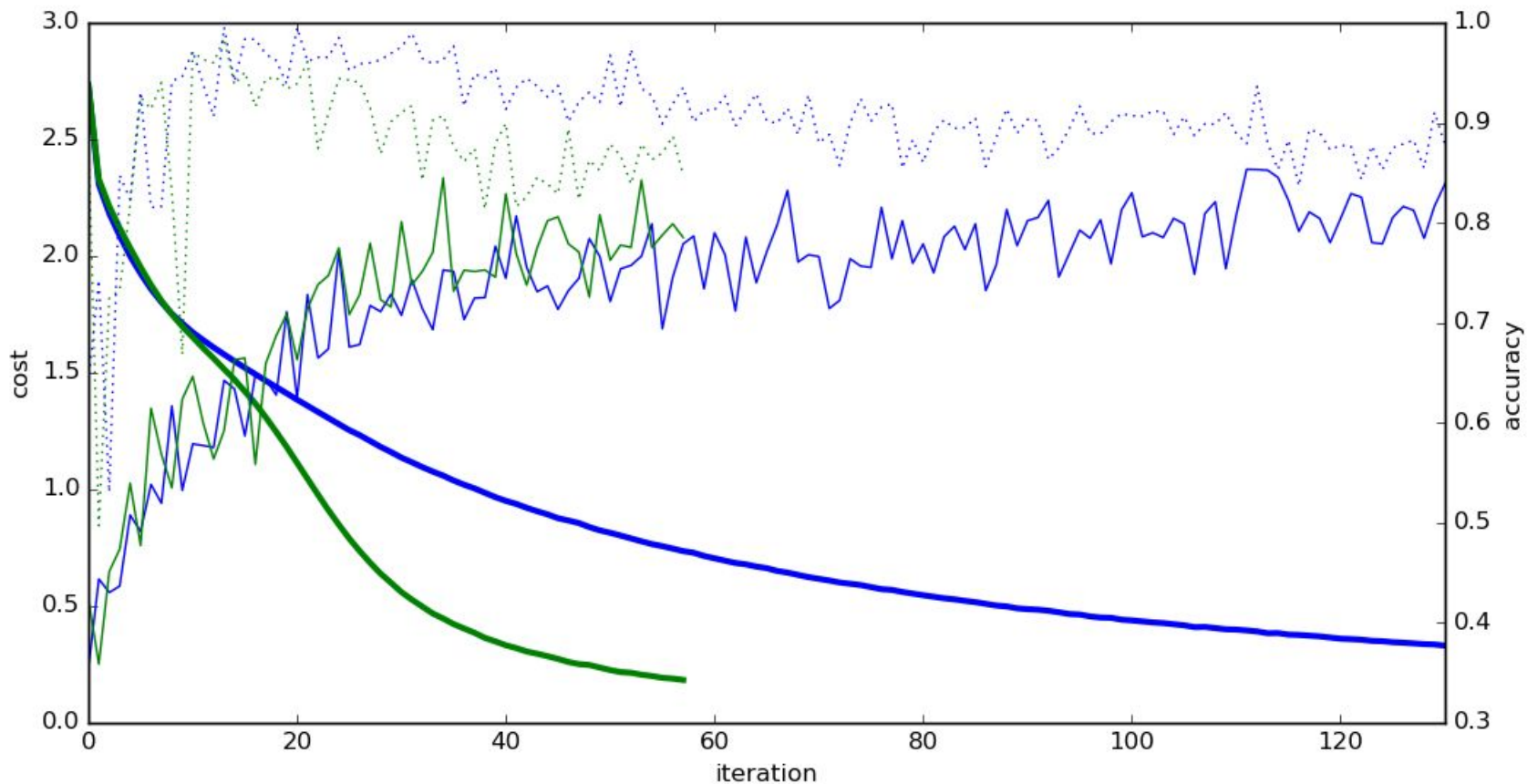
# Measuring accuracy



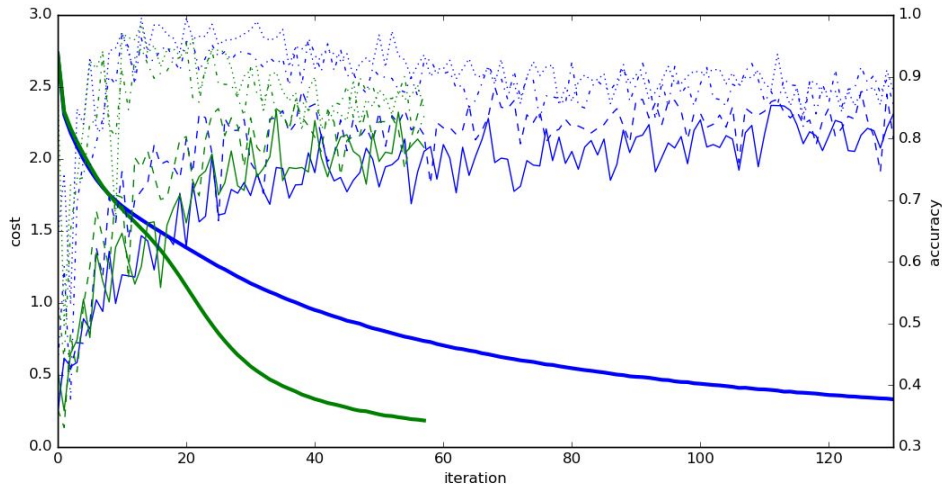
# Neurons



# Neurons: 512 vs 1024



# Neurons: 512 vs 1024



512

Low temp:

mees oli ta vastanud, et ta oli ka enam mingit neist temasse viinalt parjada sellele peale mõneda kohale ja arusaatust taastasse saati. ja see kasvab veel tema pea sees.

[...]

nii et ma teadsin, et mulle tundus mulle ka enam midagi ja ette vastama.

ma ei ole kui tema keegi meele.

High temp:

kinni üle metsavendi tema silmis peaaegu palamamaatlus. ainult ette vaaduks tõgima!, viga minu omida... esimesel tegul. meil oli ruttanud, päris kilduv paesa vananud ja just naguti olemuspoolt ühne sedasama vajaks.

1024

High temp:

ma olin karini armendama näonud paidagi elanud.

selgus, et veel siis naiselik.

tegelikult oli hetke teeline piri, oli ma auendaid kordi alla peen, seda tuli ta rihumad loodus.

# Examples

kuid minul on lõpuks ta võimalust kuidagi õhtuvalt sisse ja siis ta kõiges mõistab.

ning seda oli teinud, ta oli see pool elanud küll.

me esimesel metsas oli ta kõrval avastuse kui kallusliku hõbemil mälestuses tema poole pikaks juhtunud vaid pidevult sinna piirada ja selle nime.

ka ta silmade ja kui kaarel peitus ja enne kui ma üleselnud kui kellega teisel pool tulevad lahti.

# Conclusions

Good initial hyperparameters for 1 MB dataset.

Lackluster text, needs more data.

Original post vs this work - Estonian more complex than English?



Thank you!

