A study on different neural network structure for eatable and poisonous mushroom classification

Big Data and Deep Learning 2018
April 16th, 2018

3117999011	PUTTAPIRAT Pargorn
3117999104	CHAN MIN WAI
3117999241	HUE Lee

***All members has equal contribution,** name order is based on student ID

This work is licensed under a Creative Commons Attribution 4.0 International License.

Overview

- Dataset
- Model
- Results & Discussion

This work is licensed under a Creative Commons Attribution 4.0 International License.

Dataset

This work is licensed under a Creative Commons Attribution 4.0 International License.

Dataset

The Mushroom Dataset
by Jeff Schlimmer

Details of the dataset

# of Sample	8124 [48.20% positive | 51.80% negative]
# of Attribute/Feature	22 [21 is used]
Classification Category	binary classification

Link: https://archive.ics.uci.edu/ml/datasets/Mushroom

This work is licensed under a Creative Commons Attribution 4.0 International License.

Dataset

The Mushroom Dataset by Jeff Schlimmer

Preprocessing

cap-shape

gill-spacing

gill-size

stalk-root

cap-surface

bruises?

ring-number

gill-color

odor

stalk-surface-below-ring

cap-color

Attribute

stalk-color-below-ring

stalk-shape

ring-type

population

gill-attachment

veil-color

stalk-color-above-ring

habitat

stalk-surface-above-ring

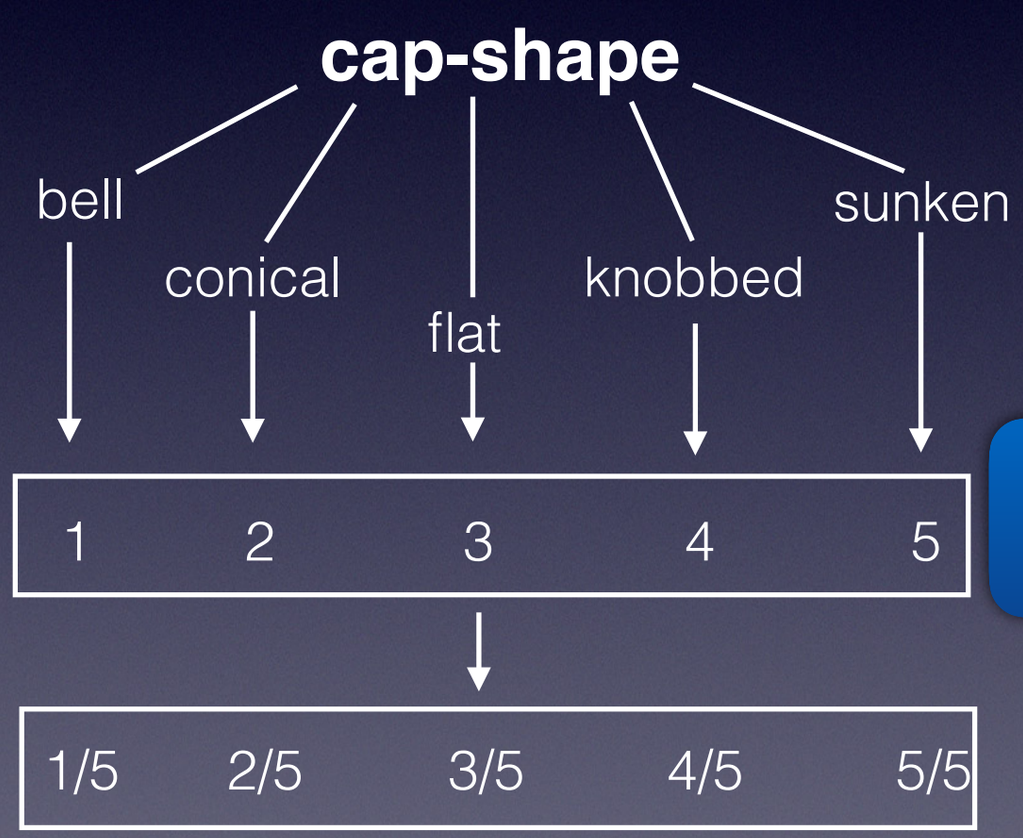
spore-print-color

This work is licensed under a Creative Commons Attribution 4.0 International License.

Dataset

The Mushroom Dataset by Jeff Schlimmer

Preprocessing



Attribute

gill-spacing

gill-size

bruises?

stalk-surface-below-ring

odor

stalk-color-below-ring

population

gill-attachment

stalk-color-above-ring

habitat

stalk-surface-above-ring

spore-print-color

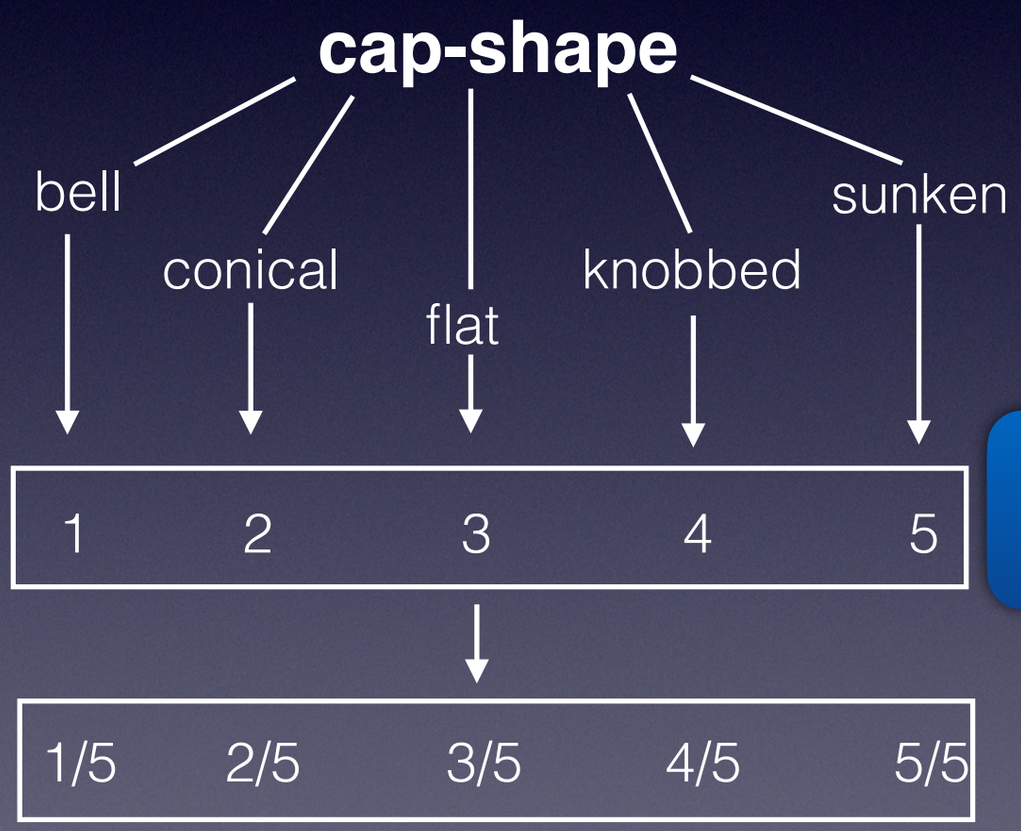
This work is licensed under a Creative Commons Attribution 4.0 International License.

Dataset

UCI Machine Learning Repository

The Mushroom Dataset by Jeff Schlimmer

Preprocessing



Attribute

stalk-color-below-ring

population

gill-attachment

habitat

stalk-color-above-ring

stalk-surface-above-ring

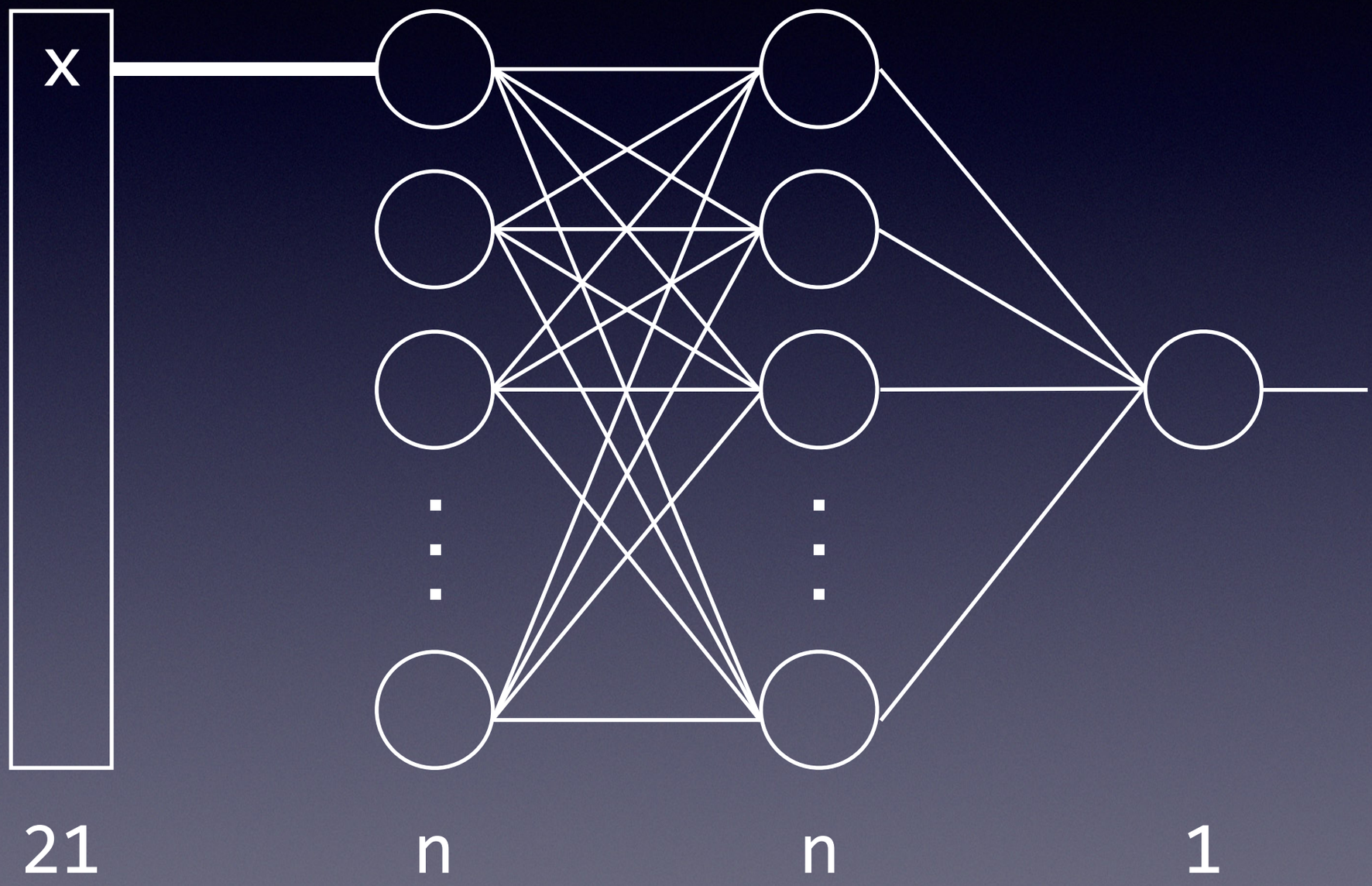
spore-print-color

This work is licensed under a Creative Commons Attribution 4.0 International License.

Network Structure

This work is licensed under a Creative Commons Attribution 4.0 International License.

Network Structure



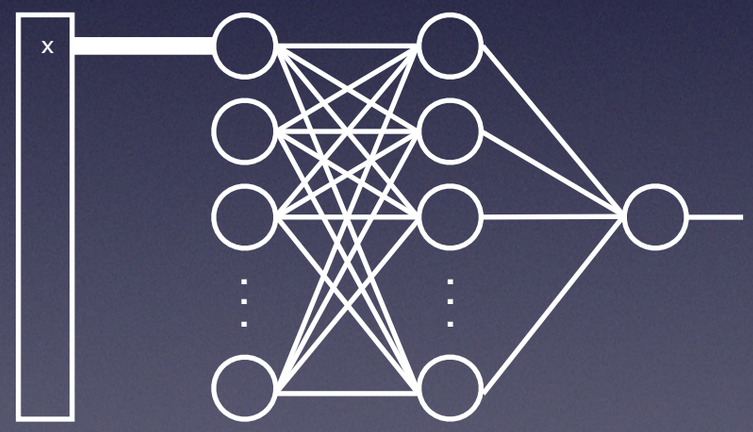
This work is licensed under a Creative Commons Attribution 4.0 International License.

Network Structure

Network type	Fully connected network
# of layers	3 [1 Input | 1 Hidden | 1 Output]

Hyperparameters

Train-Test	80% / 20%
Batch size	25 samples
Iterations	<3000
Learning Rate	0.01
Optimizer	Adamax
Loss function	MSE



Activation function: tanh()

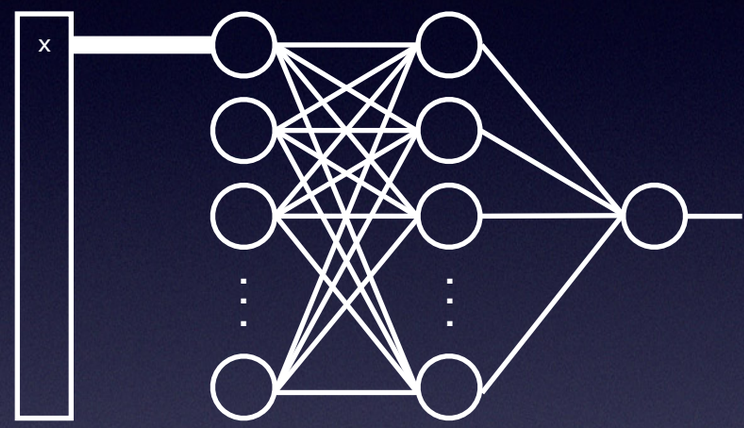
$$y = f(Wx + b)$$

This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

Training and Testing

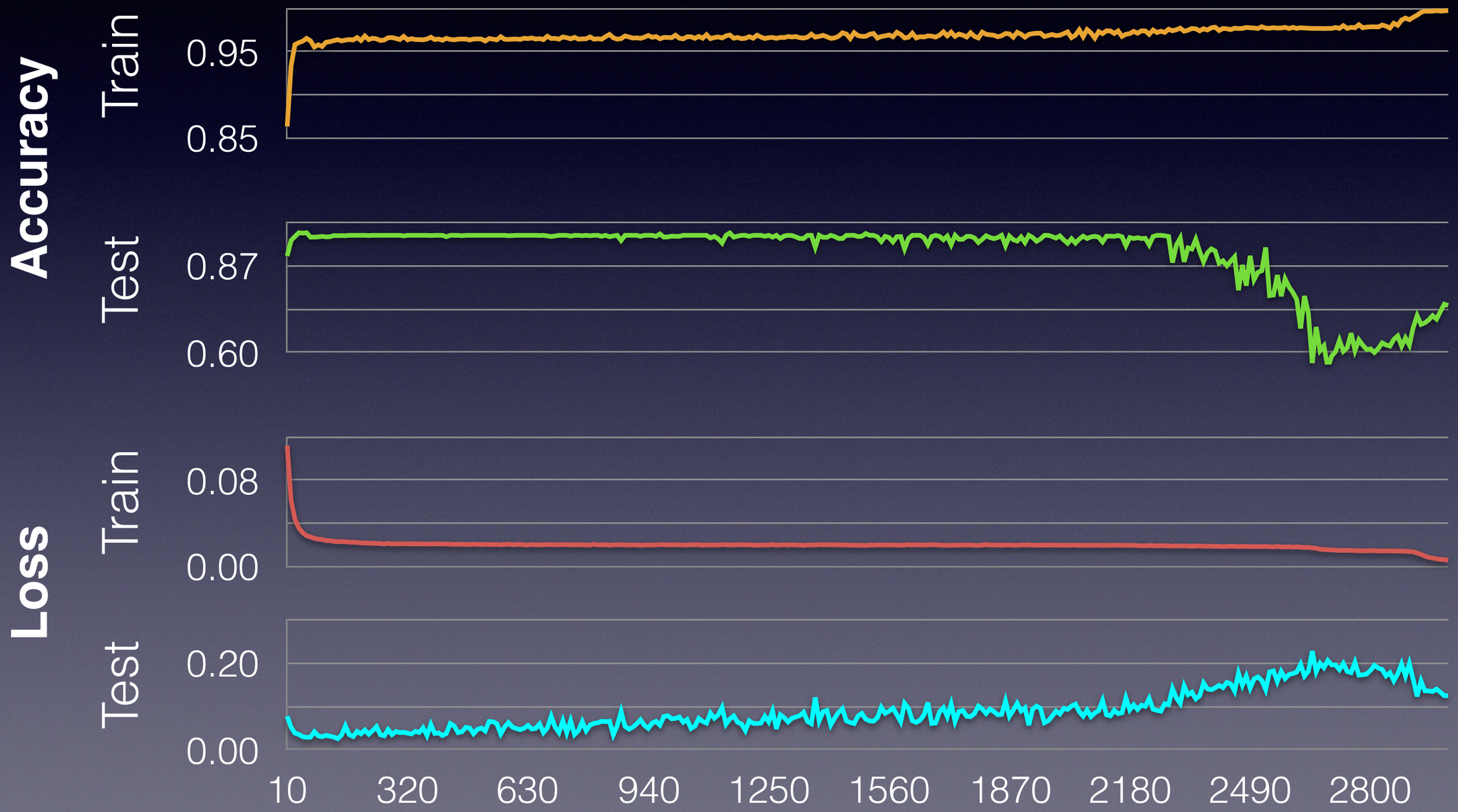
	Training Set	Test Set
Accuracy	100%	100%
Error	0%	0%



*number of nodes ≥3

This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

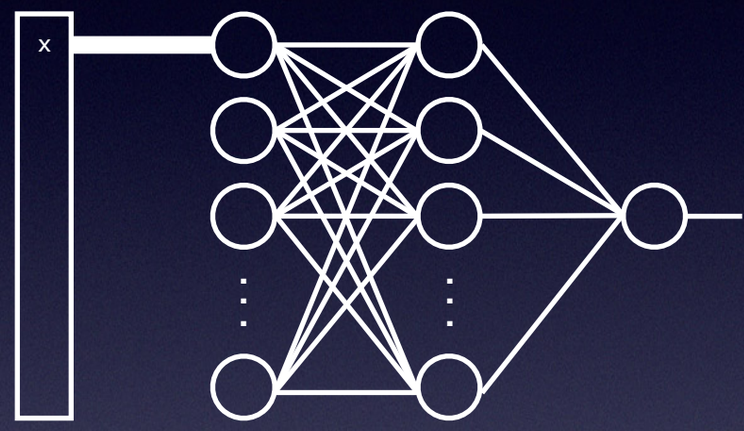


This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

Training and Testing

	Training Set	Test Set
Accuracy	100%	100%
Error	0%	0%



*number of nodes ≥3

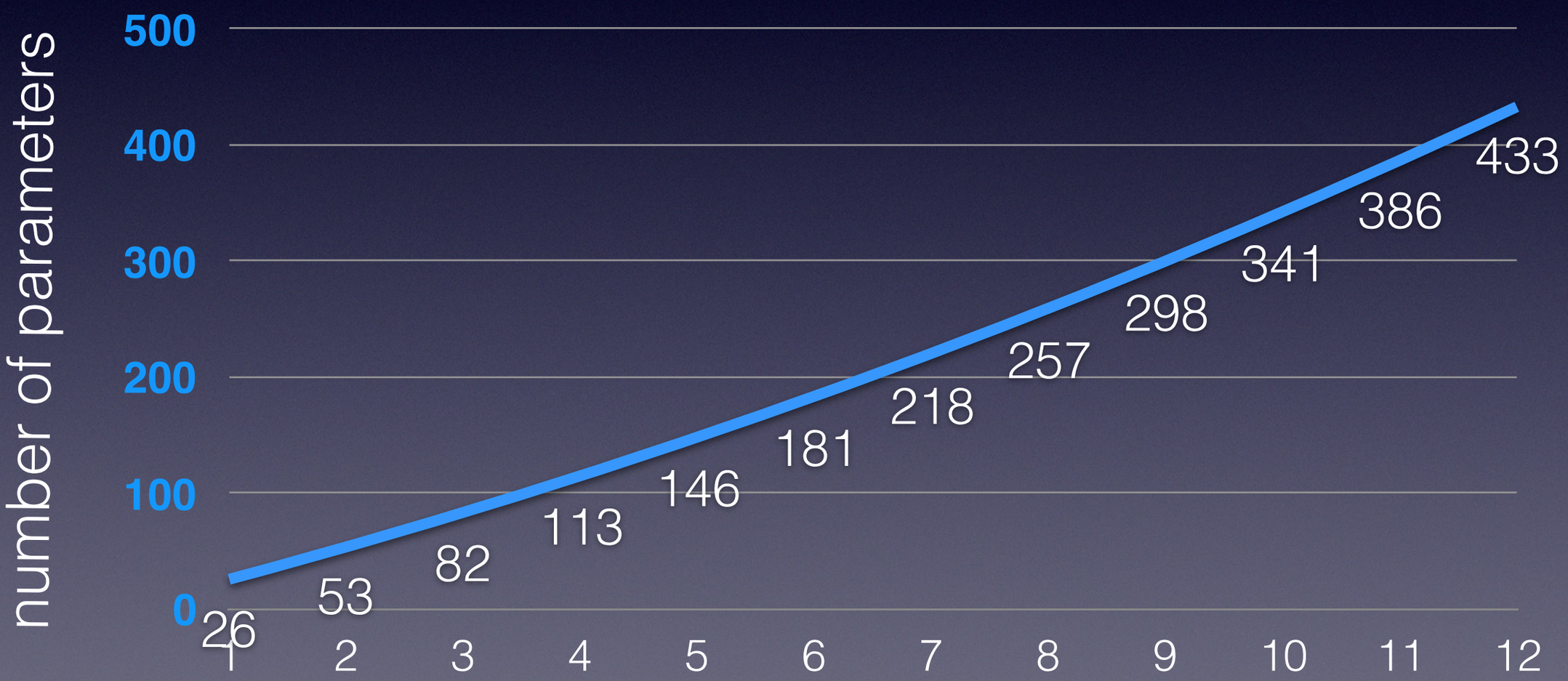
Which model is the best?

- Number of **parameters**
- Speed of **convergence**

This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

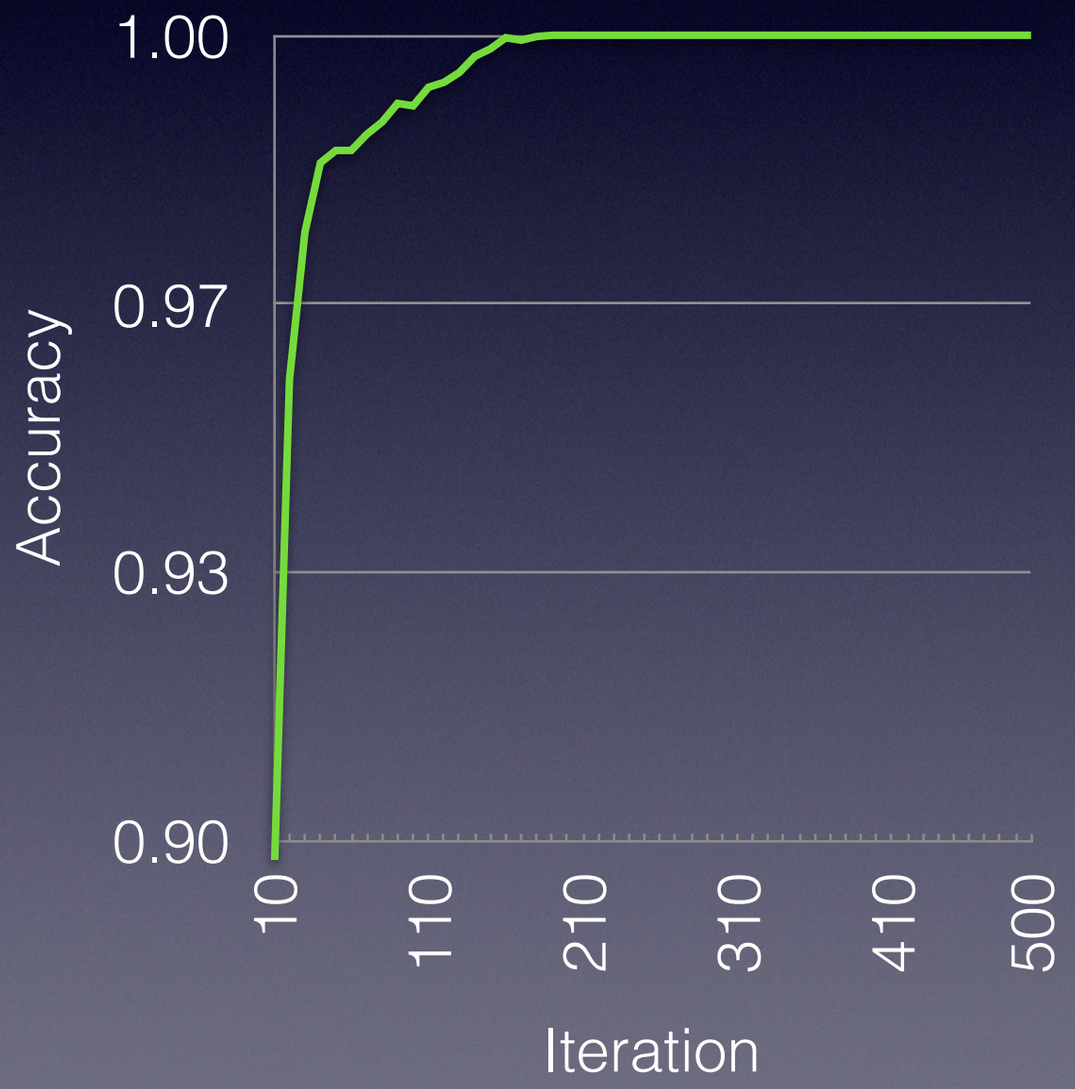
Performance of the network

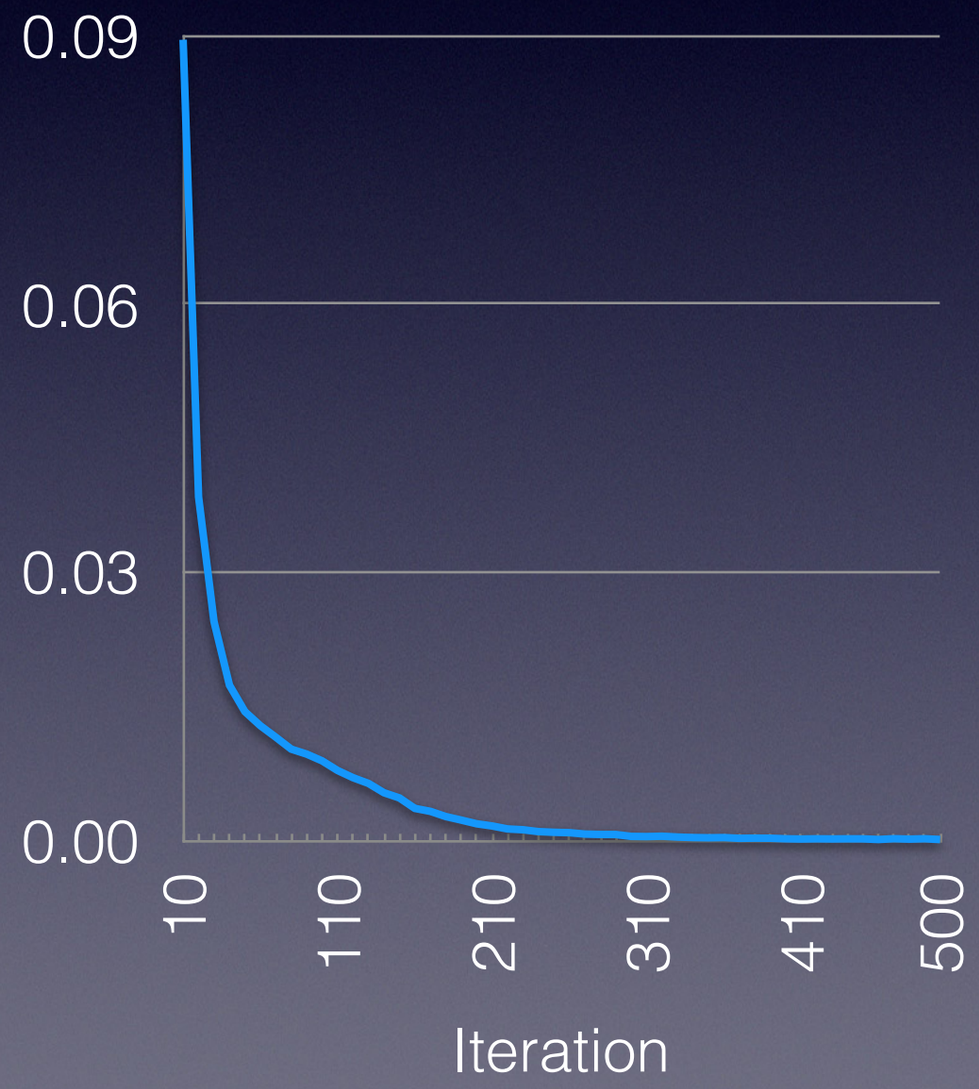


This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

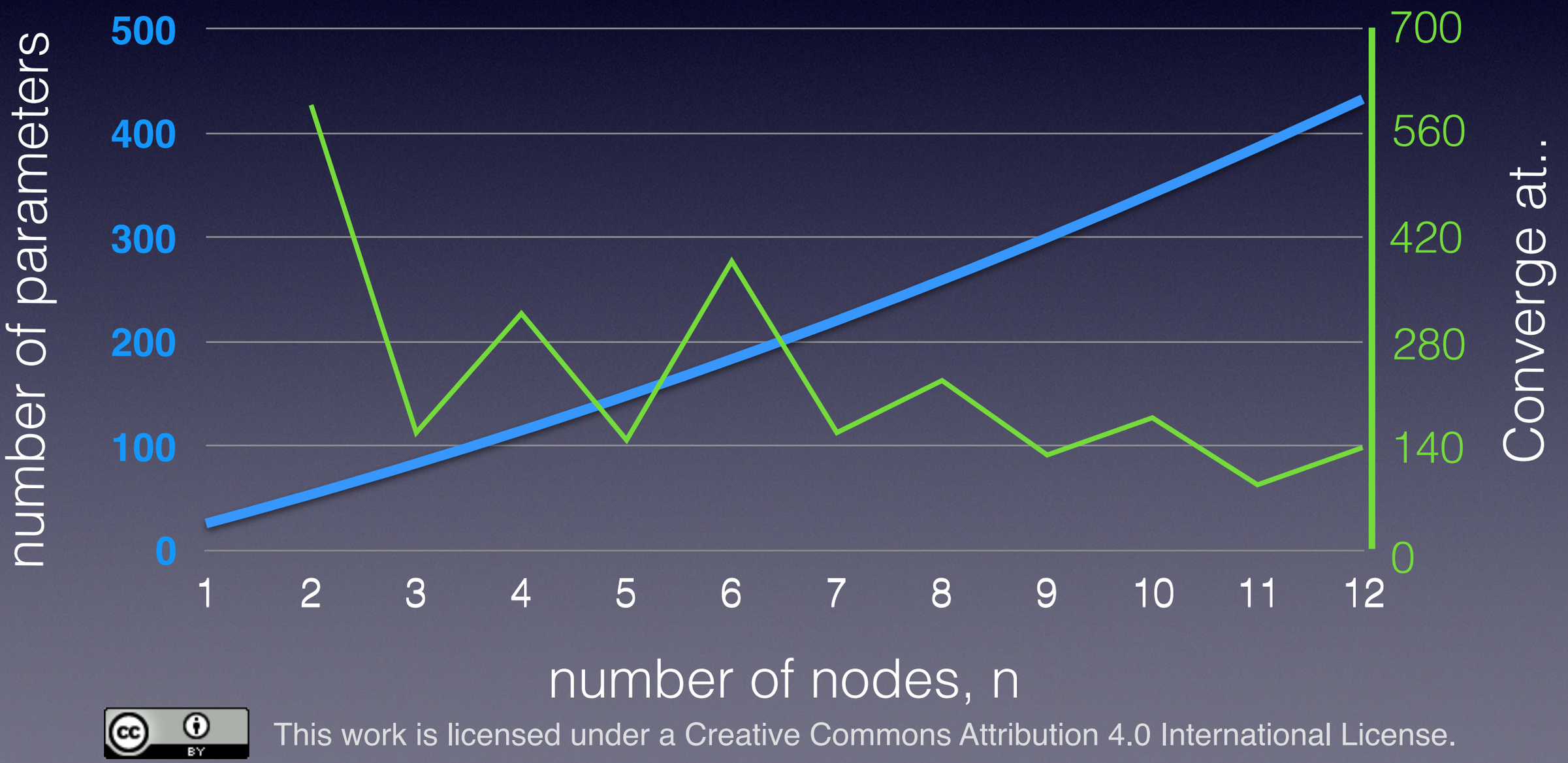
Usual training accuracy and loss [10 10 1]





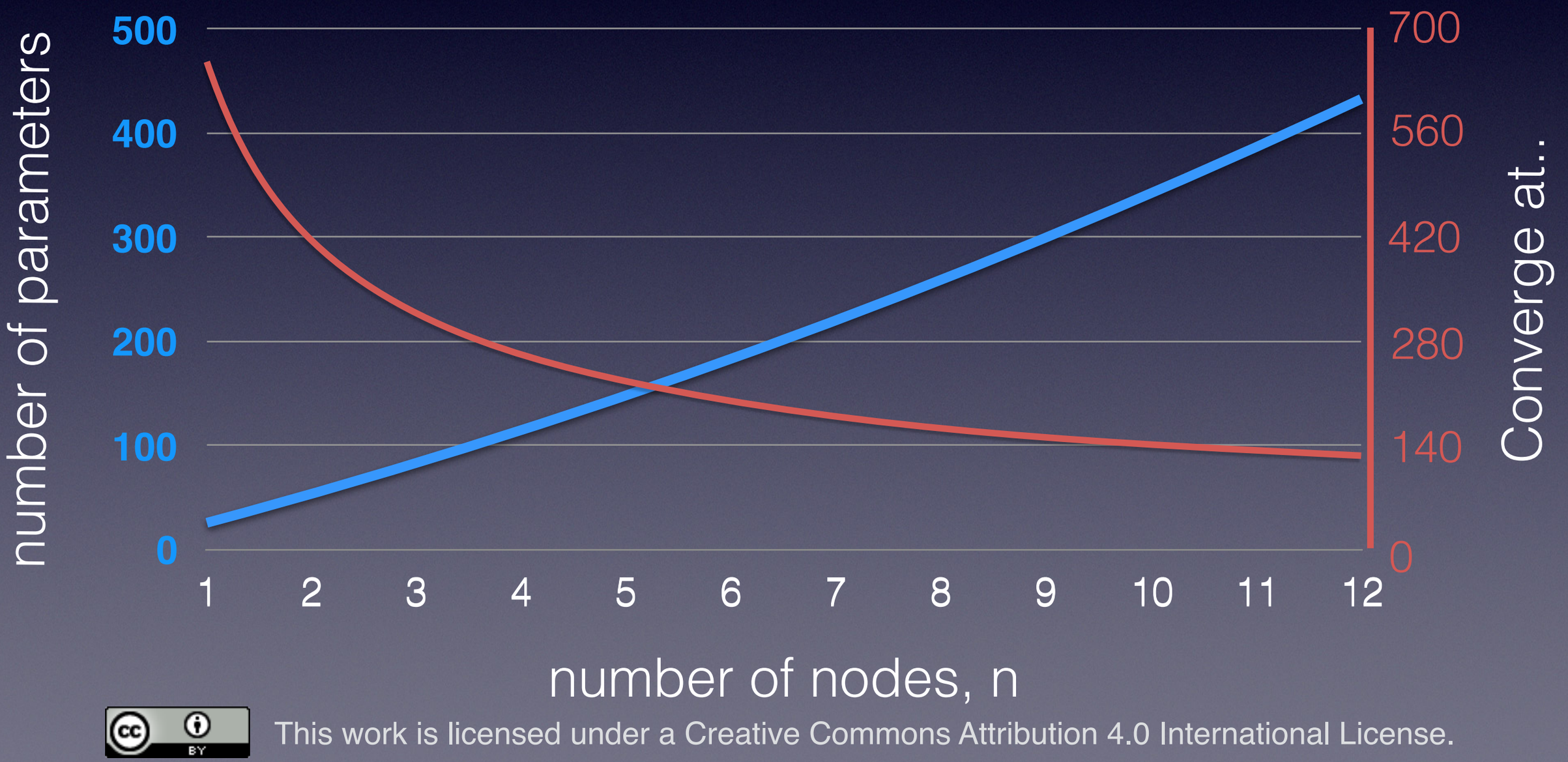
This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions



This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions



This work is licensed under a Creative Commons Attribution 4.0 International License.

Results and Discussions

- The simplest network structure that can generalize is [3 3 1] (82 parameters).
  - Overfitting can be seen in [2 2 1].

- Reducing number of features
  can reduce complexity of the network.

- Some features can be transformed into one-hot feature vector.

This work is licensed under a Creative Commons Attribution 4.0 International License.

Thanks ...