

INTRODUCTION TO DEEP LEARNING

*by Ray Cai (LC85640)
from PMC, ISG, ICG*

AGENDA

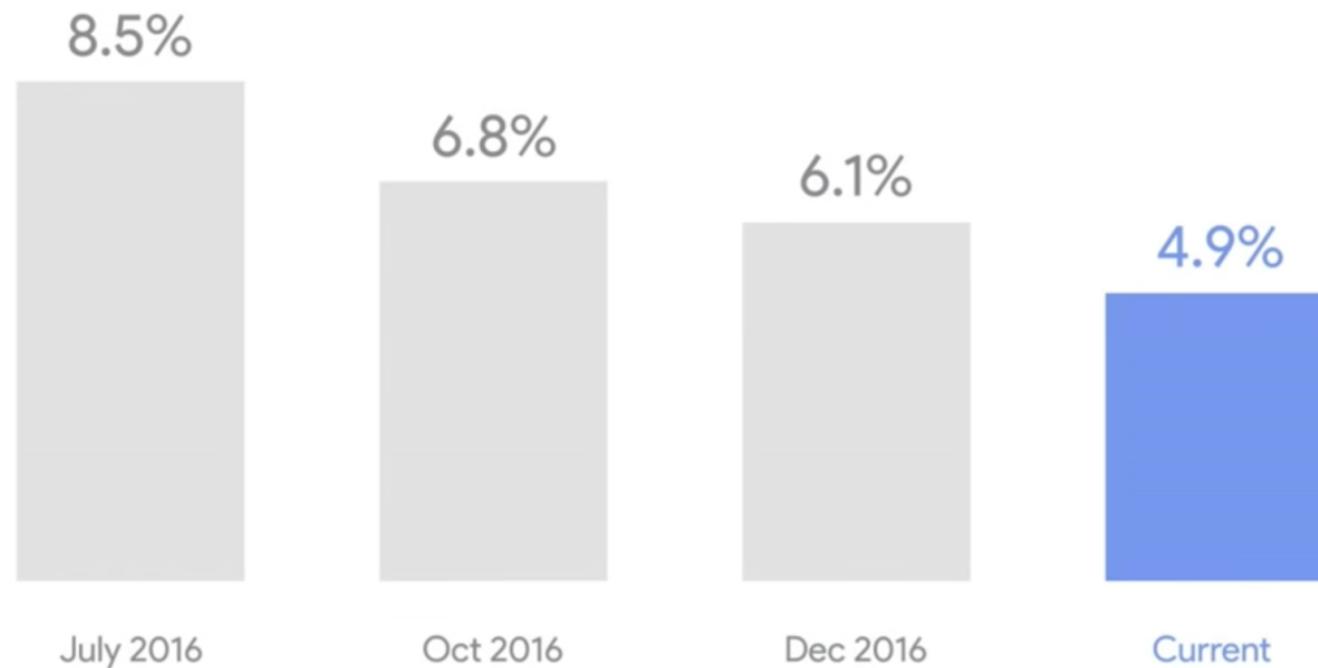
1. Application Areas
2. Approaches

APPLICATION AREAS

- Speech Recognition
- Image Recognition
- Natural Language Processing
- Visual Art Processing
- Bioinformatics

SPEECH RECOGNITION

Speech Recognition Word Error Rate



US English only.

- Google Speech API

SPEECH RECOGNITION

CLOUD SPEECH API PRICING

Powerful Speech Recognition

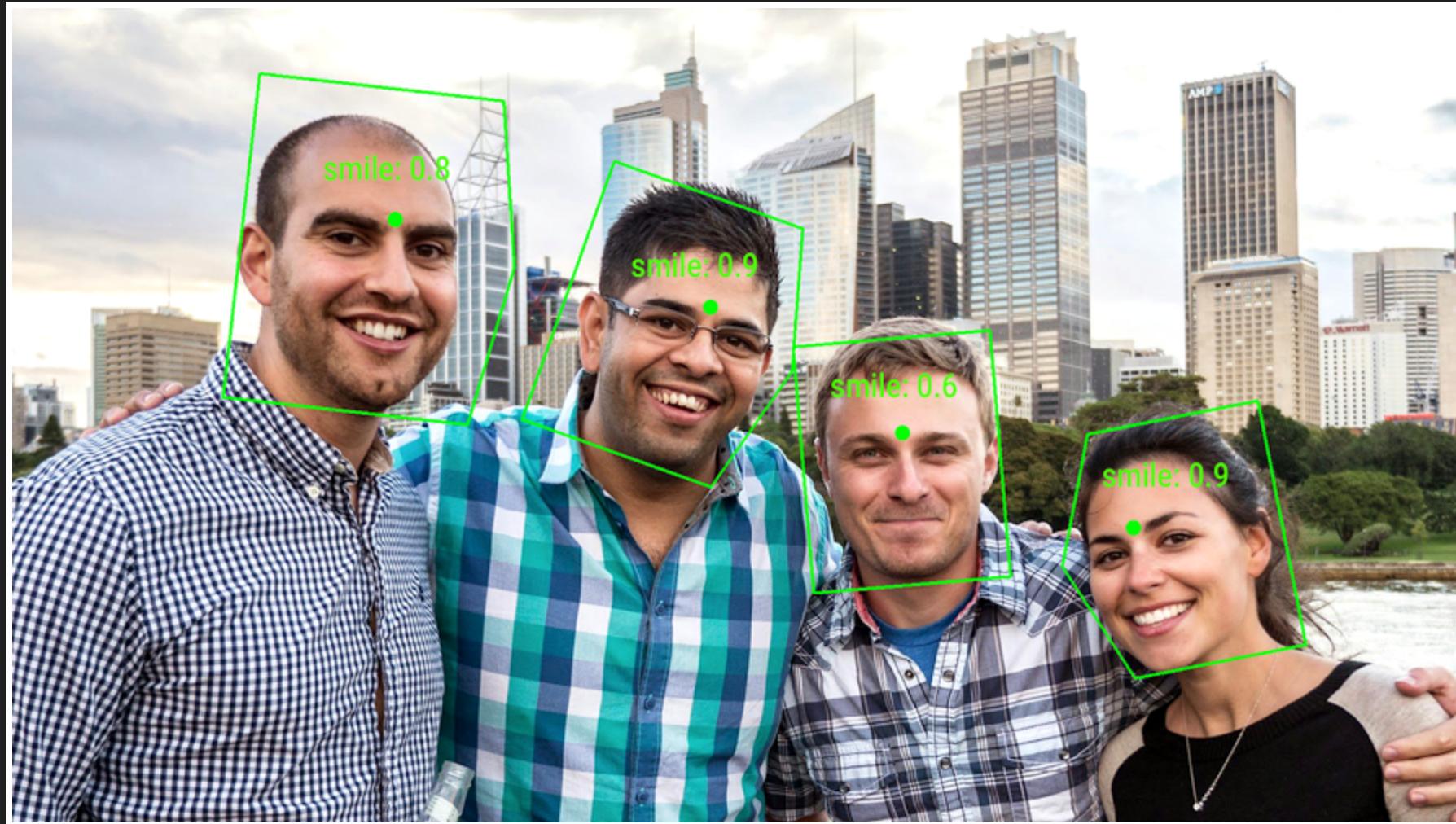
Cloud Speech API is priced per 15 seconds of audio processed after a 60 minute free tier. For details, please see our [pricing guide](#).

MONTHLY USAGE	PRICE PER 15 SECONDS*
0 - 60 minutes	Free
61 - 1,000,000 minutes**	\$0.006

\$1.44 per Hour

- [Google Speech API](#)

IMAGE RECOGNITION



- Detect Faces and Emotions

IMAGE RECOGNITION



"running", "score": 0.99803412,
"marathon", "score": 0.99482006

- Google Cloud Vision API enters Beta, open to all to try!

IMAGE RECOGNITION



{"joyLikelihood": "VERY_LIKELY"}

{"description": "ABIERTO\\n",
"local": "es"}

- Google Cloud Vision API enters Beta, open to all to try!

NATURAL LANGUAGE PROCESSING

Try the API

Citi's mission is to serve as a trusted partner to our clients by responsibly providing financial services that enable growth and economic progress. Our core activities are safeguarding assets, lending money, making payments and accessing the capital markets on behalf of our clients. We have 200 years of experience helping our clients meet the world's toughest challenges and embrace its greatest opportunities. We are Citi, the global bank – an institution connecting millions of people across hundreds of countries and cities.

[See supported languages](#)

ANALYZE

Entities Sentiment Syntax Categories

<Citi>₁'s <mission>₄ is to serve as a trusted <partner>₅ to our <clients>₂ by responsibly providing financial <services>₃ that enable <growth>₆ and economic <progress>₇. Our core <activities>₁₄ are safeguarding <assets>₁₃, lending <money>₁₂, making <payments>₉ and accessing the <capital markets>₁₁ on behalf of our <clients>₁₀. We have 200 years of <experience>₈ helping our <clients>₂ meet the <world>₂₁'s toughest <challenges>₂₀ and embrace its greatest <opportunities>₂₂. We are <Citi>₁, the global <bank>₁ – an <institution>₁ connecting <millions>₁₆ of <people>₁₇ across <hundreds>₁₈ of <countries>₁₉ and <cities>₁₅.

1. Citi	ORGANIZATION
Sentiment: Score 0 Magnitude 0	
Wikipedia Article	
Salience: 0.37	
2. clients	PERSON
Sentiment: Score 0 Magnitude 0.1	
Salience: 0.29	
3. services	OTHER
Sentiment: Score -0.1 Magnitude 0.3	
Salience: 0.08	
4. mission	OTHER
Sentiment: Score 0 Magnitude 0	
Wikipedia Article	

- Google Natural Language API

NATURAL LANGUAGE PROCESSING

Try the API

Citi's mission is to serve as a trusted partner to our clients by responsibly providing financial services that enable growth and economic progress. Our core activities are safeguarding assets, lending money, making payments and accessing the capital markets on behalf of our clients. We have 200 years of experience helping our clients meet the world's toughest challenges and embrace its greatest opportunities. We are Citi, the global bank – an institution connecting millions of people across hundreds of countries and cities.

[See supported languages](#)

ANALYZE

Entities Sentiment Syntax Categories

Dependency Parse Label Part of Speech Lemma Morphology

pos
Citi
NOUN
number=SINGULAR
proper=PROPER

ps
's
PRT

nsubj
mission
NOUN
number=SINGULAR

root
is
be
VERB
mood=INDICATIVE
number=SINGULAR
person=THIRD
tense=PRESENT

aux
to
PRP

xcomp
serve
VERB

prep
as
ADP

det
a
DET

amod
trusted
ADJ
tense=PAST

pobj
partner
NOUN
number=SINGULAR

prep
to
ADP

poss
our
PRON
case=GENITIVE
number=PLURAL
person=FIRST

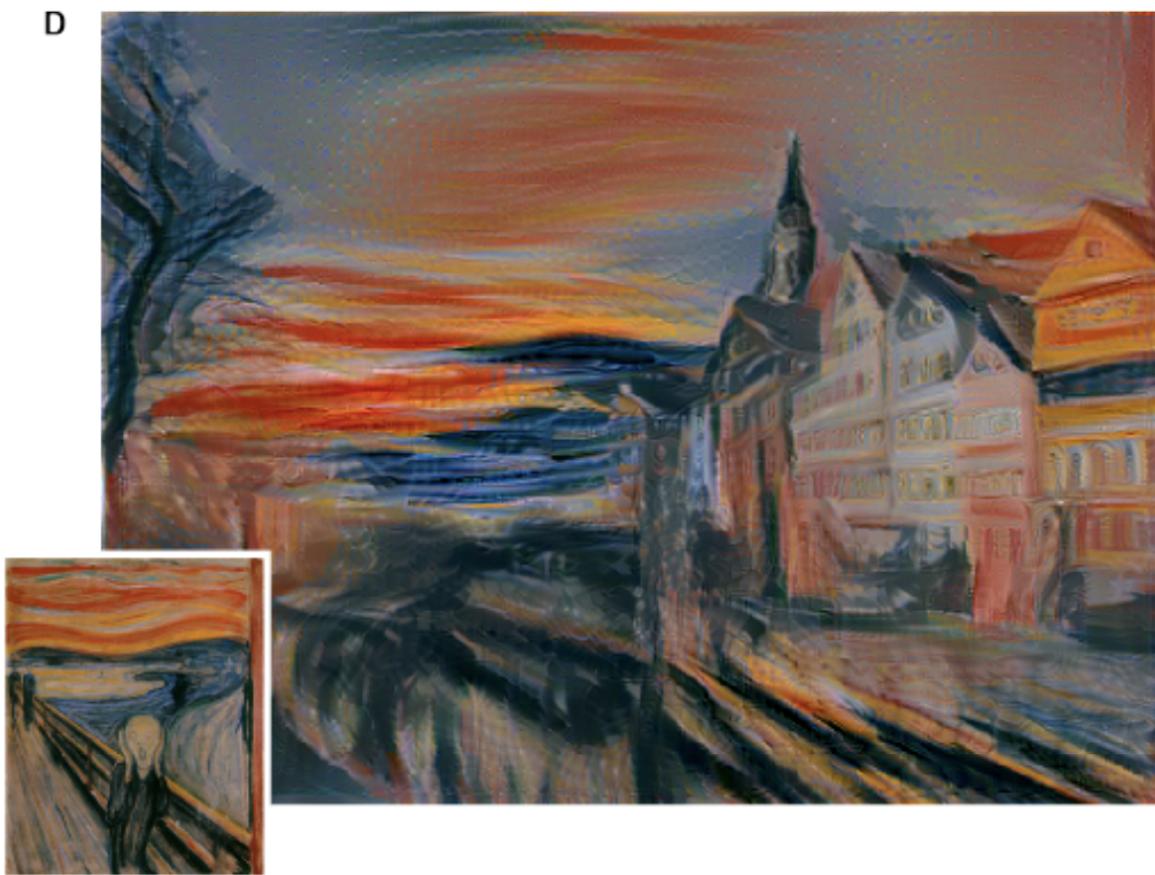
- Google Natural Language API

VISUAL ART PROCESSING



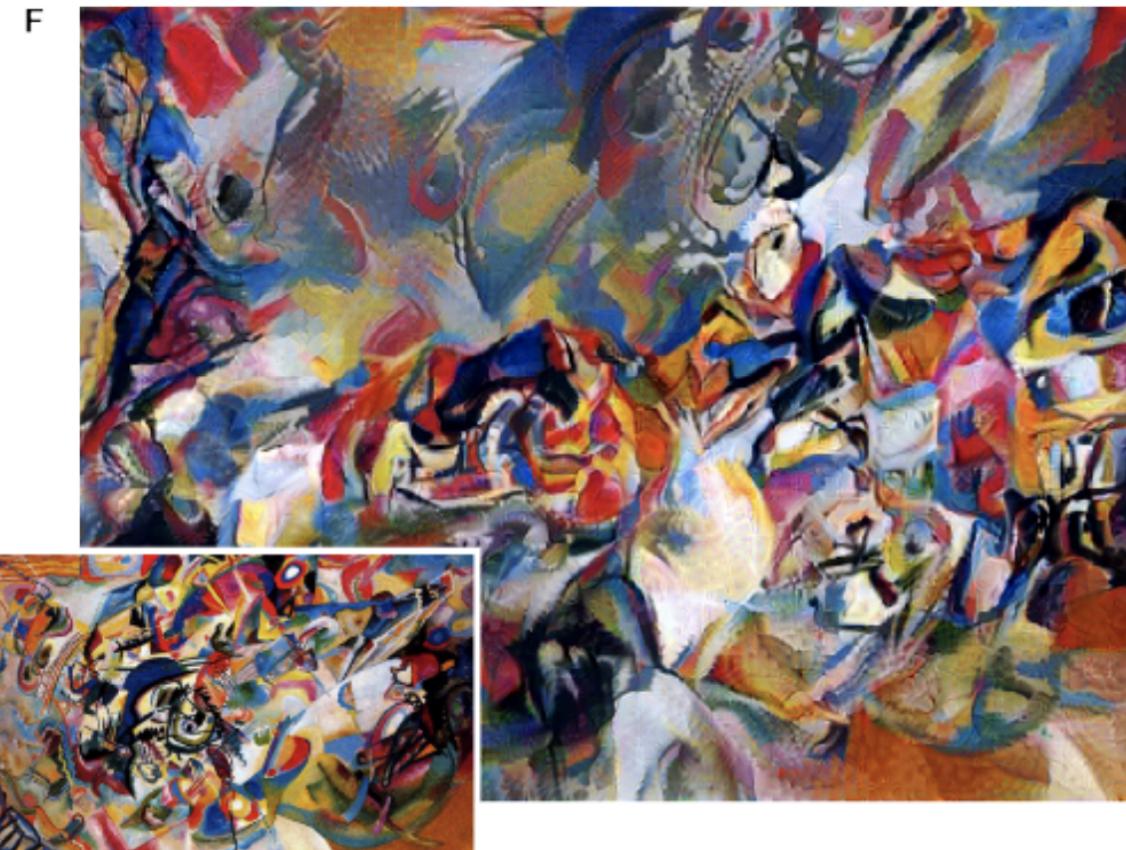
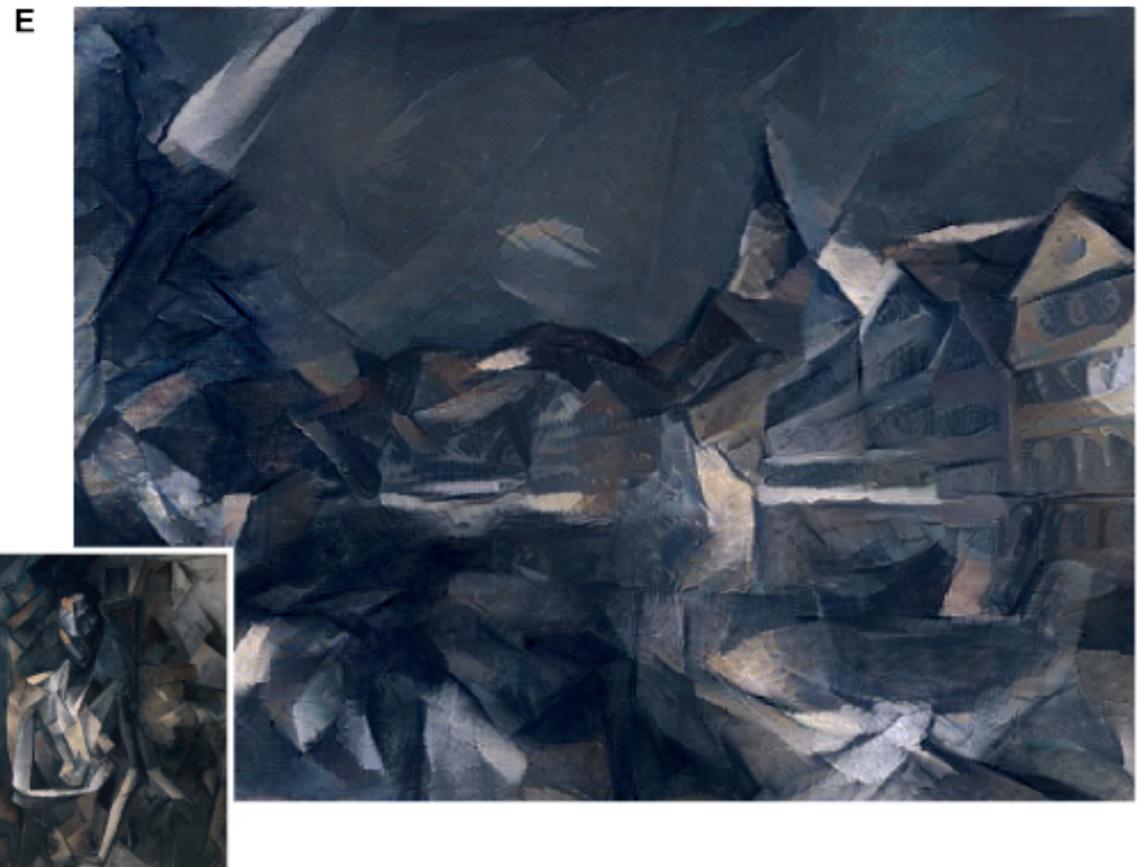
- A Neural Algorithm of Artistic Style

VISUAL ART PROCESSING



- A Neural Algorithm of Artistic Style

VISUAL ART PROCESSING



- A Neural Algorithm of Artistic Style

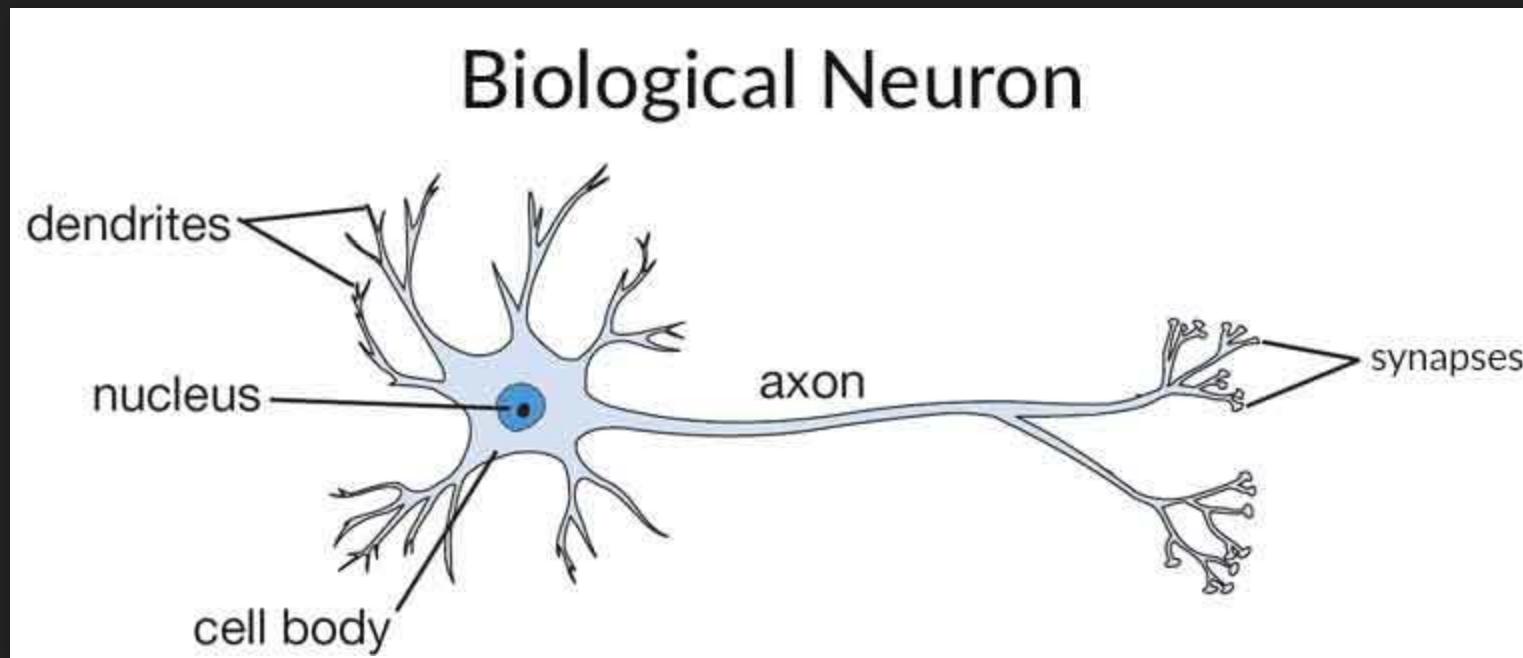
APPROACHES

DEEP LEARNING

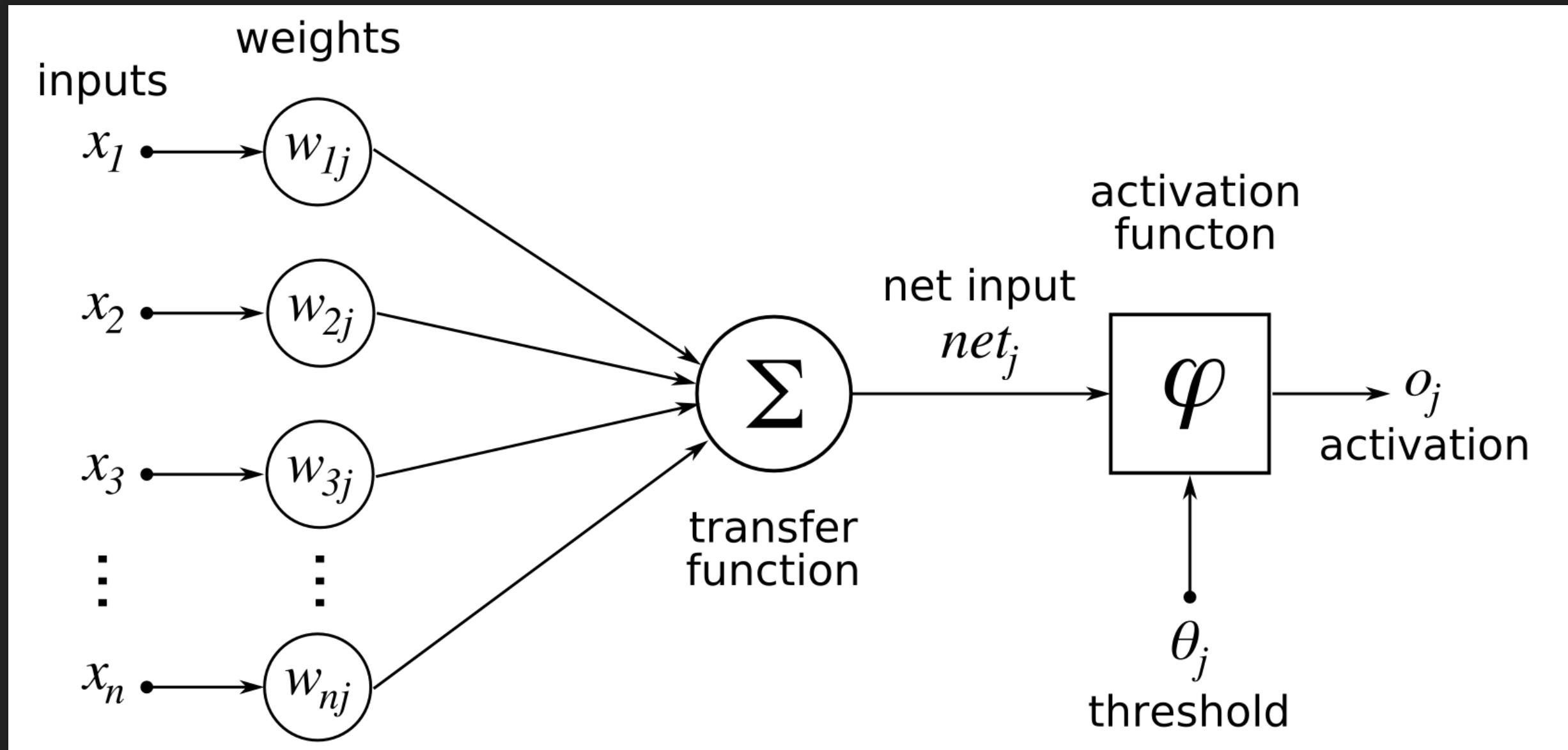
- Machine Learning
 - Support Vector Machine
 - Artificial Neural Network
 - **Deep learning**
 - Decision Tree
 - ...

ARTIFICIAL NEURAL NETWORK

BIOLOGICAL NEURON

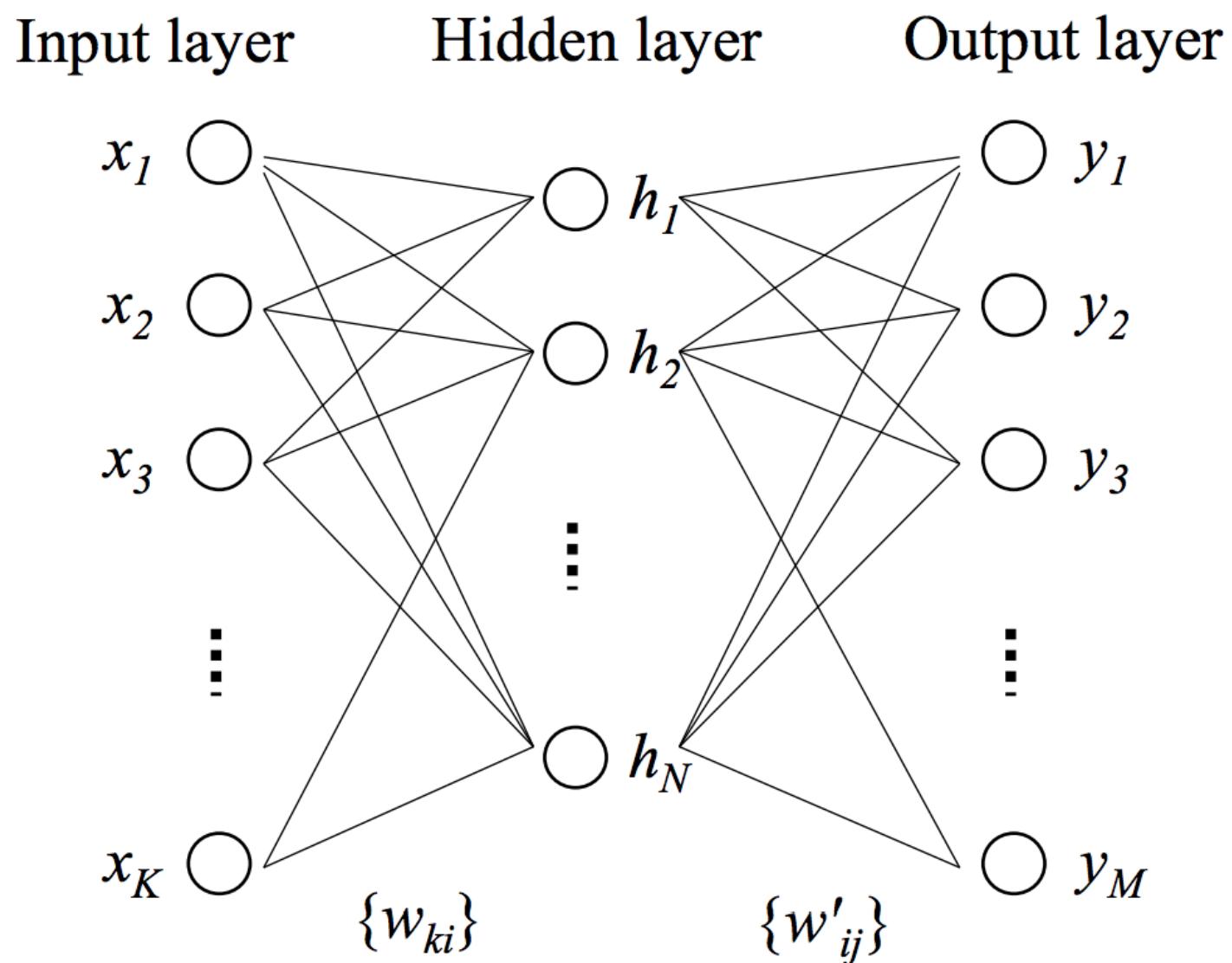


MATHEMATICAL NEURON

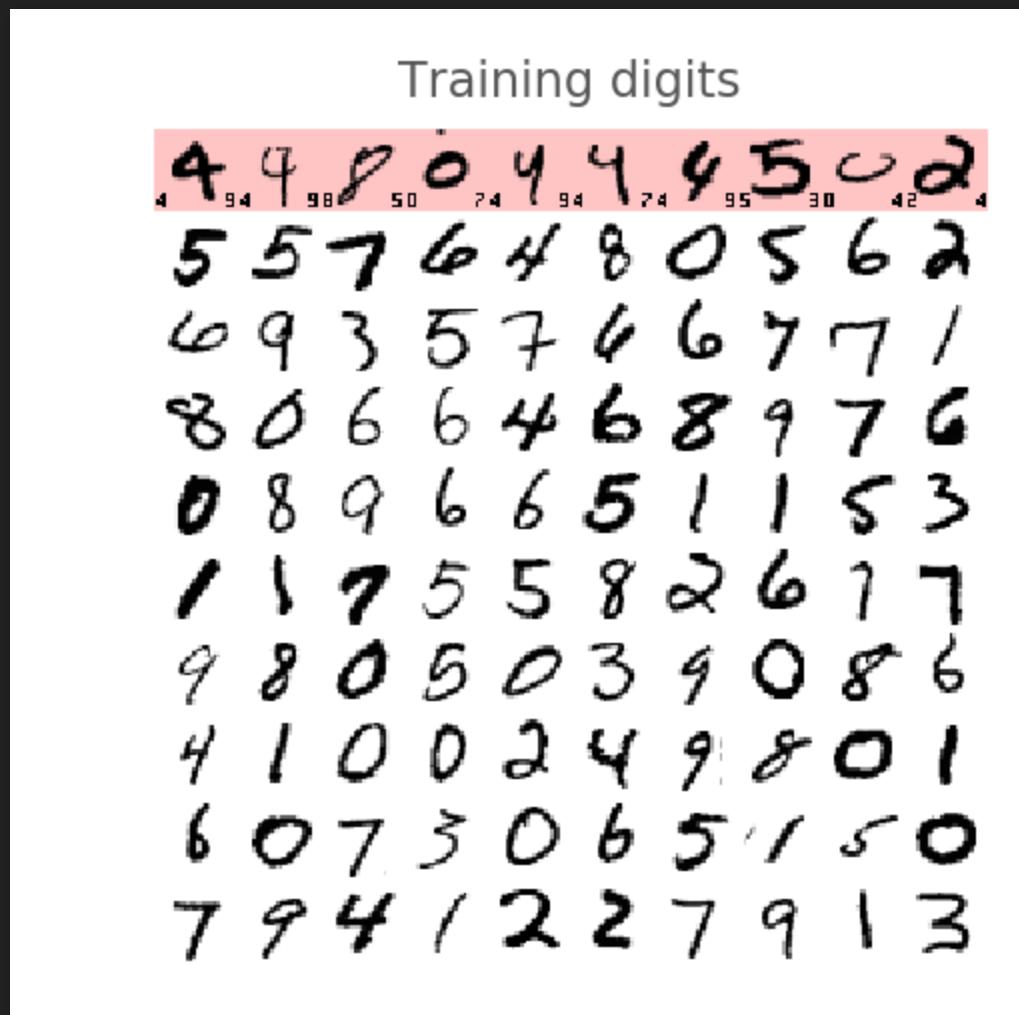


$$f = \psi\left(\sum(X \times W + b)\right)$$

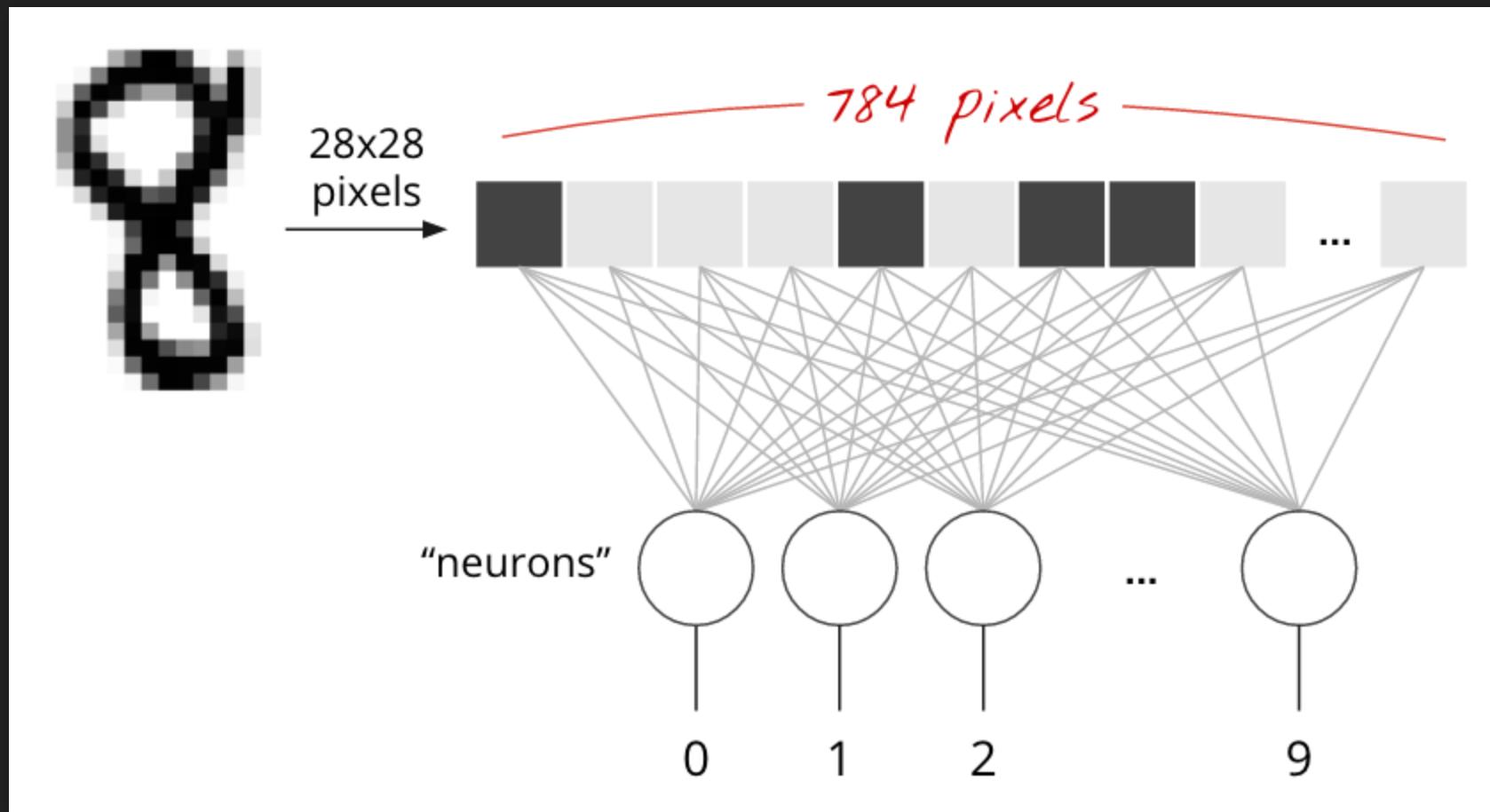
NEURAL NETWORK



HANDWRITTEN DIGITS CLASSIFICATION



1-LAYER NEURAL NETWORK



1-LAYER NEURAL NETWORK - WEIGHT

$$W = \begin{bmatrix} w_{0,0} & w_{0,1} & w_{0,2} & w_{0,3} & \dots & w_{0,9} \\ w_{1,0} & w_{1,1} & w_{1,2} & w_{1,3} & \dots & w_{1,9} \\ w_{2,0} & w_{2,1} & w_{2,2} & w_{2,3} & \dots & w_{2,9} \\ w_{3,0} & w_{3,1} & w_{3,2} & w_{3,3} & \dots & w_{3,9} \\ \dots & & & & & \\ w_{783,0} & w_{783,1} & w_{783,2} & w_{783,3} & \dots & w_{783,9} \end{bmatrix}$$

1-LAYER NEURAL NETWORK - ACTIVATION FUNCTION

$$\psi = \text{softmax}(L_n) = \frac{e^{L_n}}{\|e^L\|}$$

1-LAYER NEURAL NETWORK - FORMULA

$$Y = \text{softmax}(X \cdot W + b)$$

Predictions $Y[100, 10]$

Images $X[100, 784]$

Weights $W[784, 10]$

Biases $b[10]$

applied line by line

matrix multiply

broadcast on all lines

tensor shapes in []

The diagram illustrates the softmax formula $Y = \text{softmax}(X \cdot W + b)$ with handwritten annotations explaining the tensor shapes and operations:

- Predictions:** $Y[100, 10]$ is connected to the softmax function.
- Images:** $X[100, 784]$ is connected to the matrix multiplication operation.
- Weights:** $W[784, 10]$ is connected to the matrix multiplication operation.
- Biases:** $b[10]$ is connected to the addition operation.
- Operations:**
 - applied line by line:** This annotation points to the Y term, indicating that the softmax function is applied line by line.
 - matrix multiply:** This annotation points to the product $X \cdot W$.
 - broadcast on all lines:** This annotation points to the addition term $+ b$, indicating that the bias vector is broadcast across all lines.
- Tensor shapes:** The text "tensor shapes in []" is at the bottom left, and "tensor shapes in []" is also written vertically next to the Y term.

1-LAYER NEURAL NETWORK - LOSS FUNCTION

0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	0	1	0	0

actual probabilities, "one-hot" encoded

Cross entropy: $-\sum Y'_i \cdot \log(Y_i)$

)
computed probabilities
this is a "6"

0.1	0.2	0.1	0.3	0.2	0.1	0.9	0.2	0.1	0.1
0	1	2	3	4	5	6	7	8	9

MATHEMATICAL PROBLEM

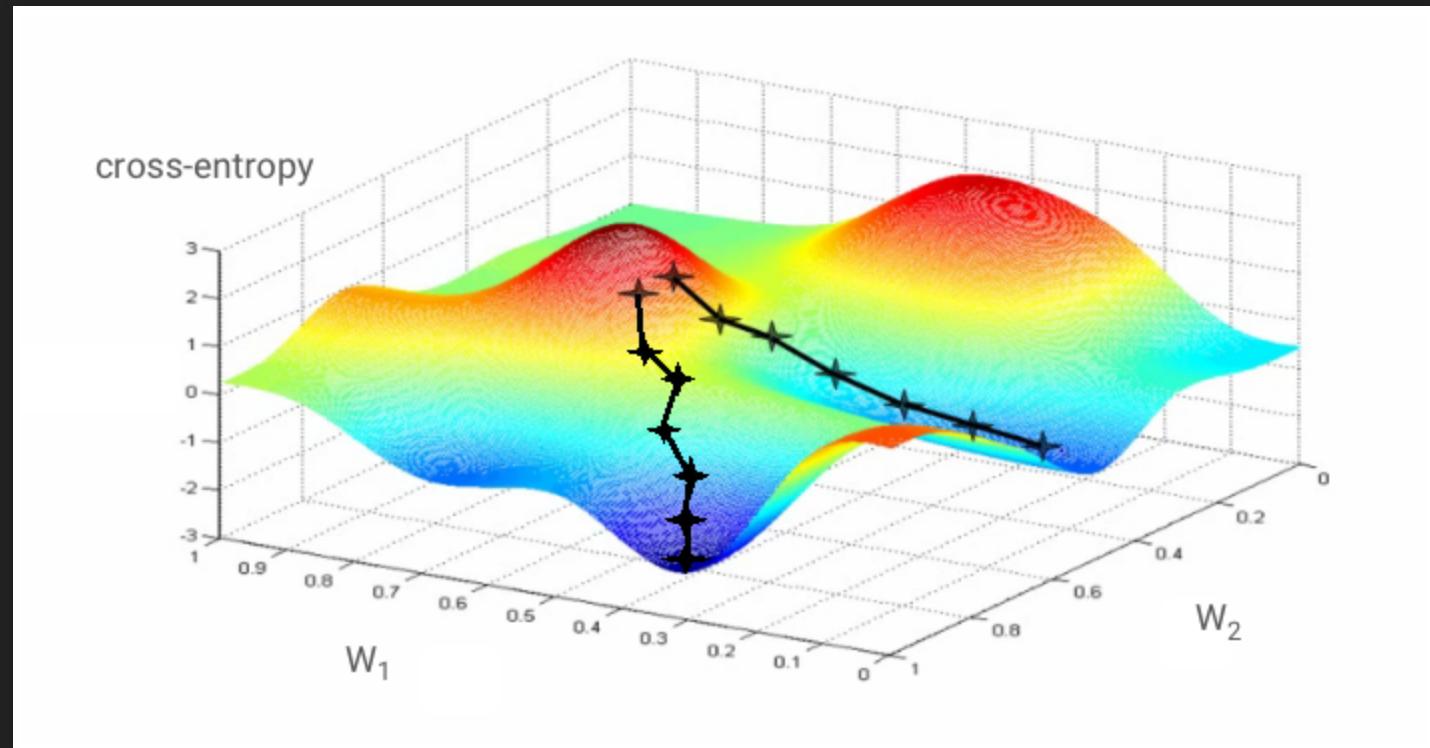
Known:

$$Y = \text{softmax}(X.W + b)$$

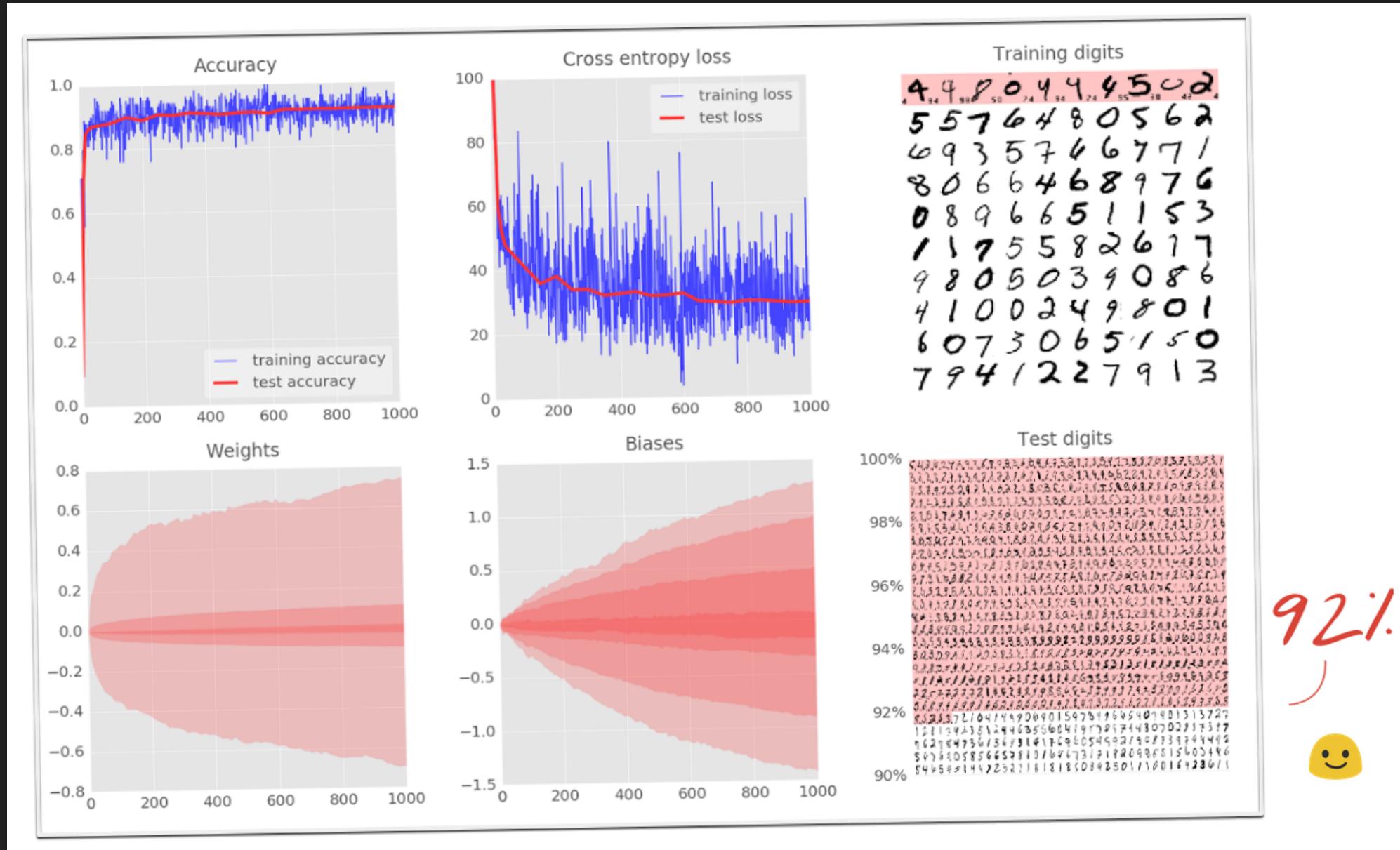
$$L = -\sum Y'_i \cdot \log(Y_i)$$

Adjust W and b minimise L

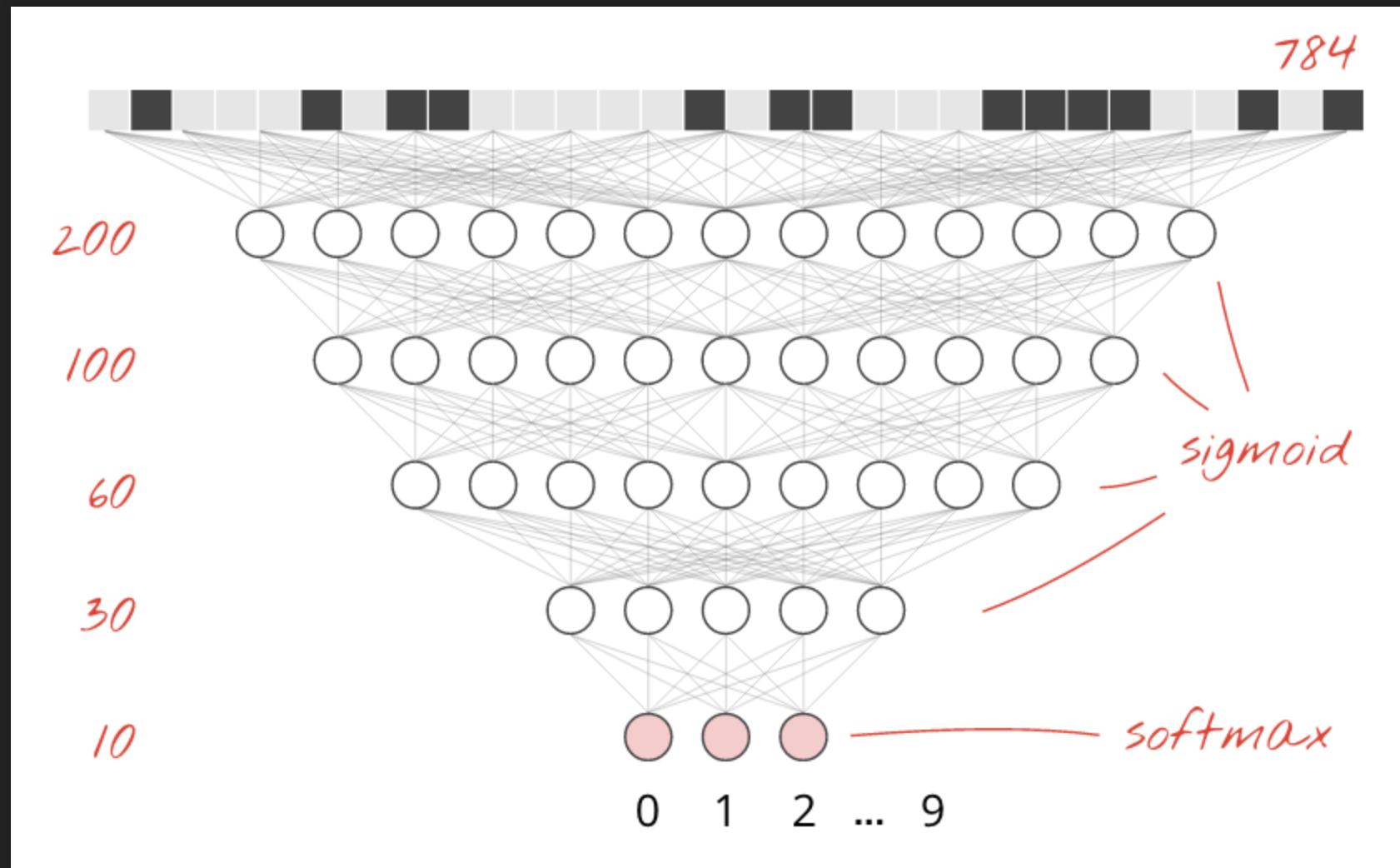
GRADIENT DESCENT



ACCURACY



DEEP NEURAL NETWORK



MATHEMATICAL PROBLEM

$$T_1 = \text{sigmoid}(X \cdot W_1 + b_1)$$

$$T_2 = \text{sigmoid}(T_1 \cdot W_2 + b_2)$$

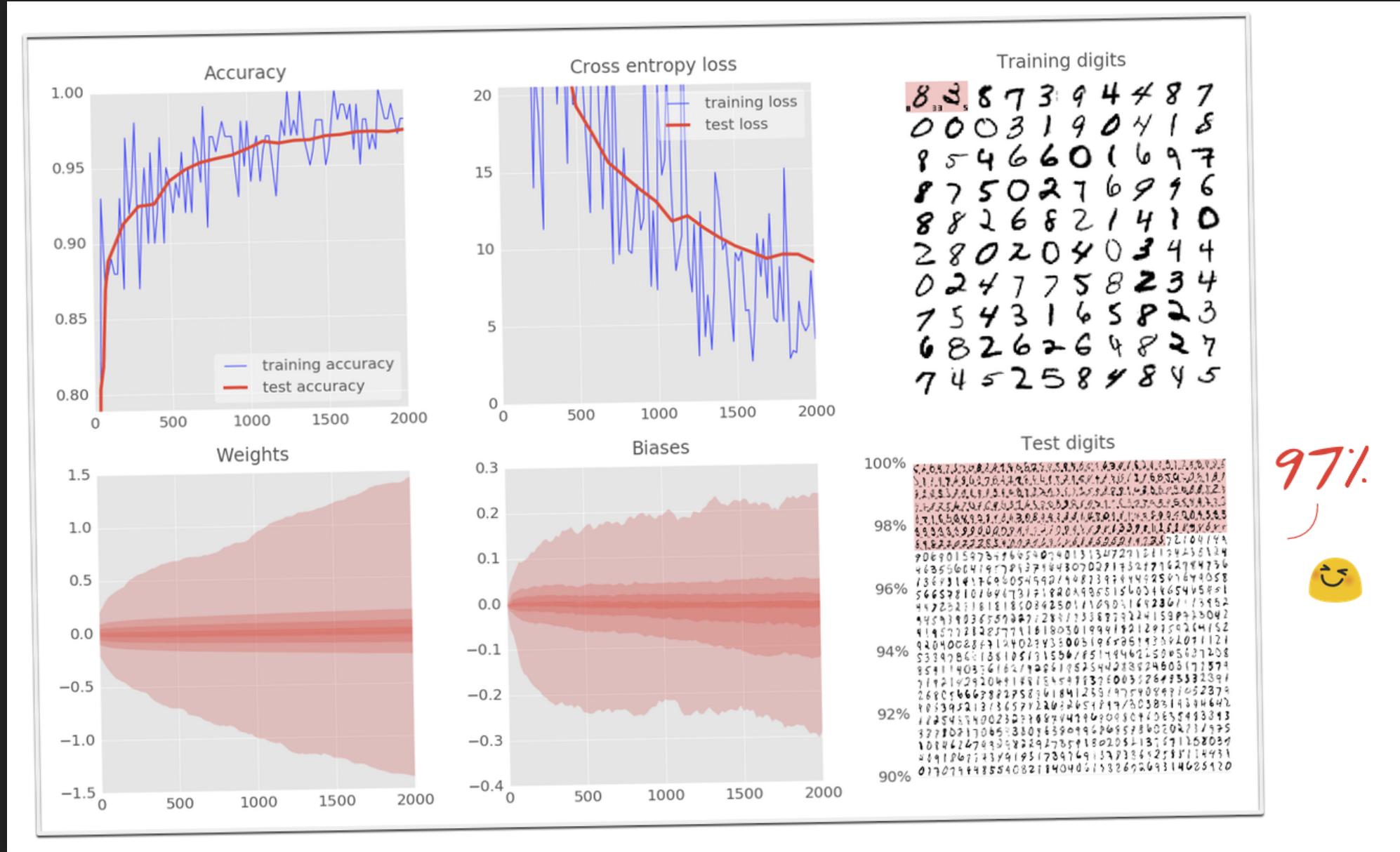
$$T_3 = \text{sigmoid}(T_2 \cdot W_3 + b_3)$$

$$T_4 = \text{sigmoid}(T_3 \cdot W_4 + b_4)$$

$$Y = \text{sigmoid}(T_4 \cdot W_4 + b_4)$$

$$L = -\sum Y'_i \cdot \log(Y_i)$$

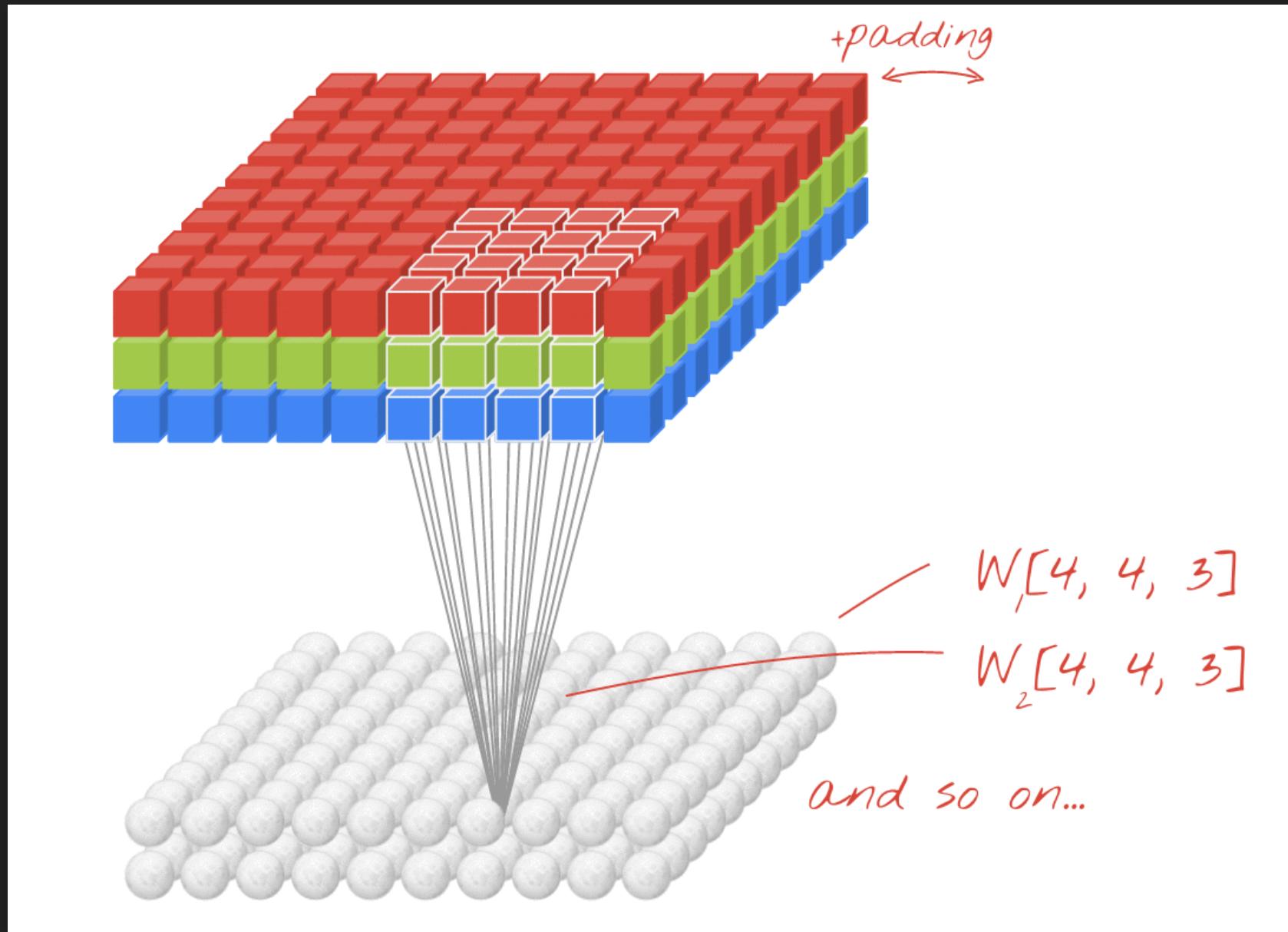
ACCURACY



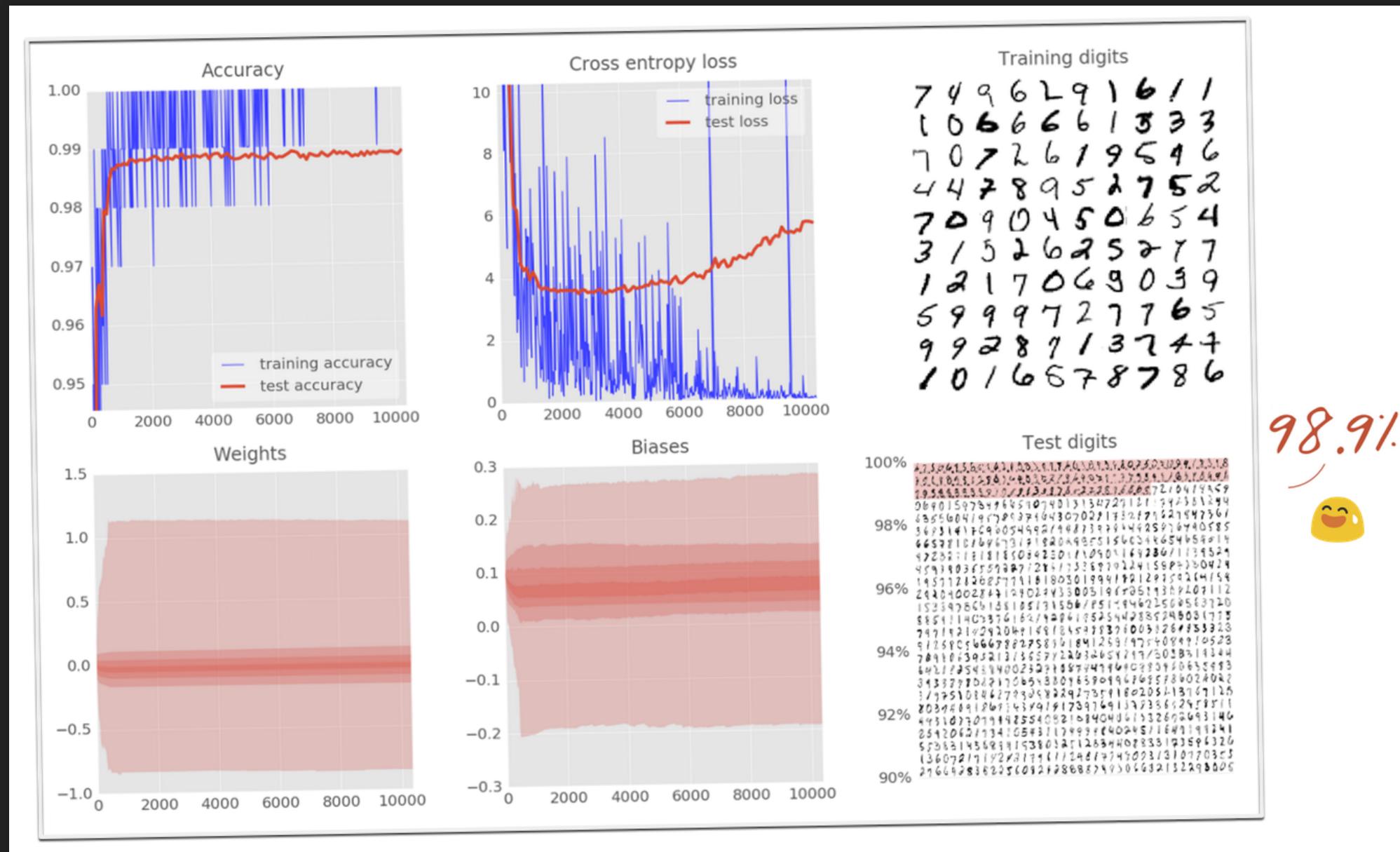
MORE COMPLEX NEURAL NETWORK

- Convolutional networks

CONVOLUTIONAL NETWORKS



ACCURACY



REFERENCE

1. TensorFlow and deep learning, without a PhD
2. Deep Learning

Q&A

THANK YOU