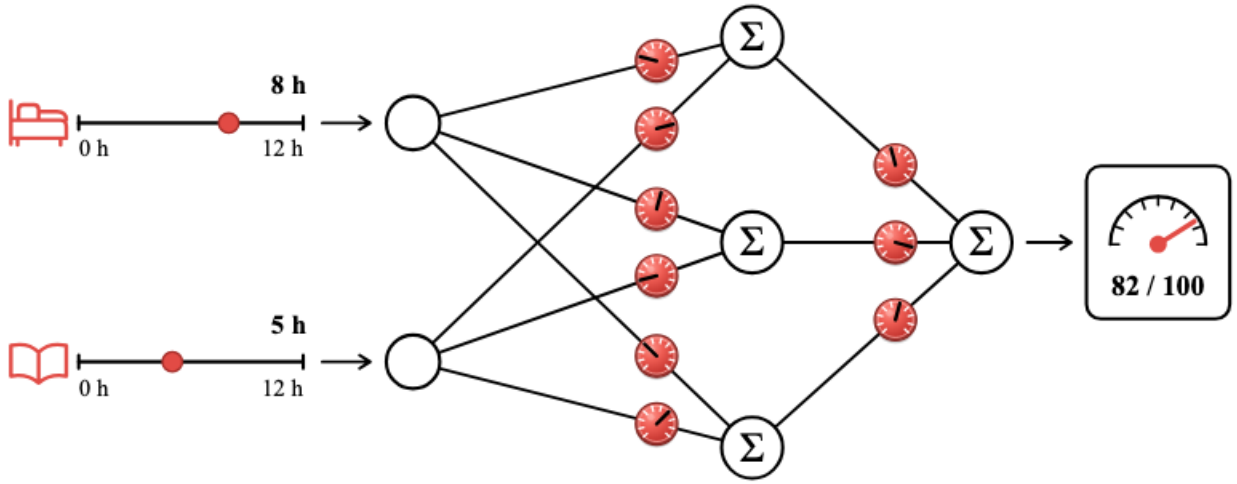


L-2

How Neural Networks Learn

AILT1001
Artificial Intelligence Literacy I

SCHOOL OF COMPUTING & DATA SCIENCE
The University of Hong Kong

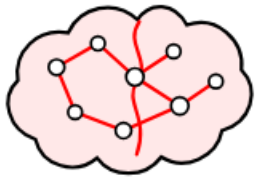
DEMO 1 · L1 Exit Ticket Results

ailt.cds.hku.hk

40 · 40 · 20 **Final Test** **Group Project** **Lecture Participation**

DEMO 2 · Recap: Course Essentials

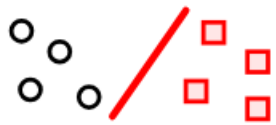
Last Week: The Path to Modern AI



Brain

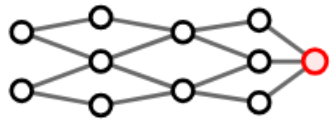
≈86 Billion Biological Neurons

Inspired



Perceptron

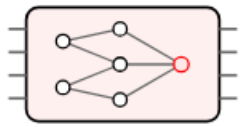
Learns Weights from Examples



Deep Learning

More Layers + Data + GPUs

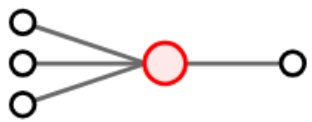
Today



Modern LLMs

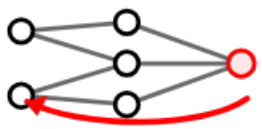
Transformers + Pretraining + Scale

1943 — 1958 — 1986 — 2012 — 2017 — 2020s



Artificial Neuron

Numbers In → Number Out



Backpropagation

Error Guides Weight Updates

Transformer

Attention Connects Context

Inspired by Biology. Built from Math, Data, and Computing.

Last Week: The Path to Modern AI

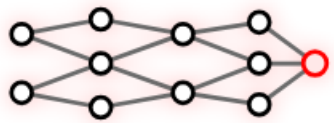
Brain

≈86 Billion Biological Neurons

Inspired

Perceptron

Learns Weights from Examples



Deep Learning

More Layers + Data + GPUs

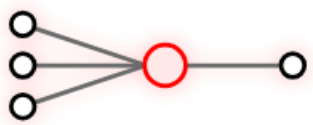
Today

Modern LLMs

Transformers + Pretraining + Scale

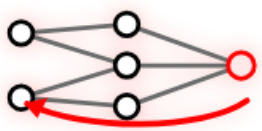
TODAY: HOW NEURAL NETWORKS LEARN

1943 — 1958 — 1986 — 2012 — 2017 — 2020s



Artificial Neuron

Numbers In → Number Out



Backpropagation

Error Guides Weight Updates

Transformer

Attention Connects Context

Inspired by Biology. Built from Math, Data, and Computing.

Why Does It Matter
How AI Learns and Makes Predictions?

Pilots Do Not Need to Know How to Build Planes

Pilots Do Not Need to Know How to Build Planes

Understanding how planes fly helps pilots recognise

CAPABILITIES · LIMITS · WARNING SIGNS

Understanding how AI learns and predicts reveals its

CAPABILITIES · LIMITS · POSSIBLE FAILURES

and helps you decide when to trust it.

What Powers These Products?

ChatGPT

Claude

DeepSeek

MINIMAX

KlingAI

TikTok

Netflix

Spotify

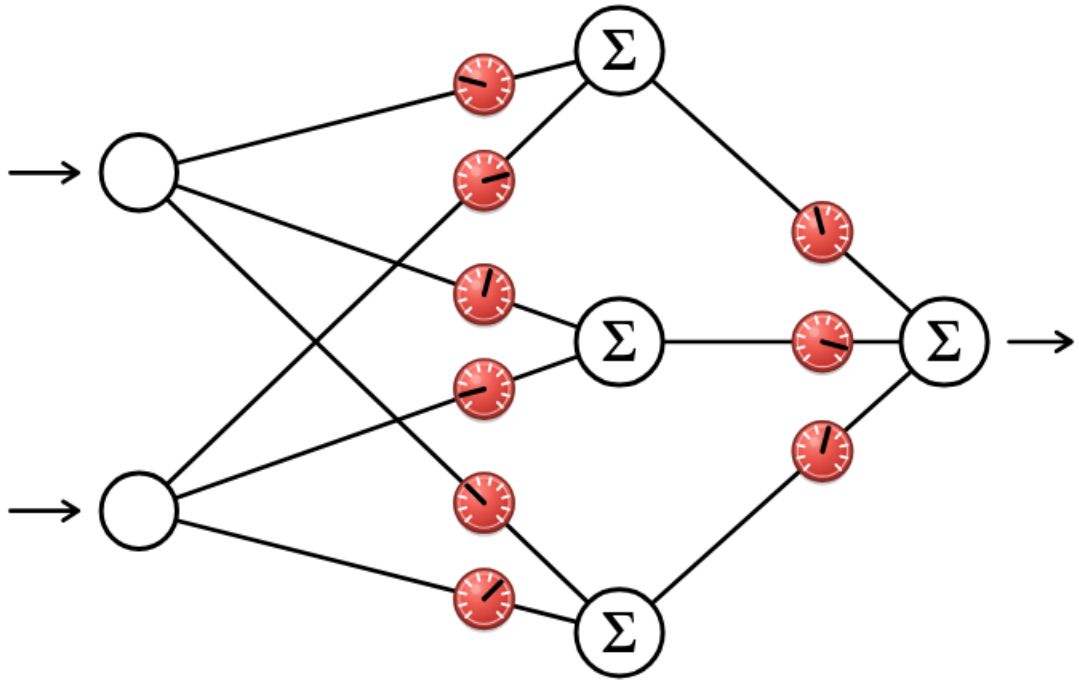
amazon

Snapchat Filters

Phone Camera

UNITREE

What Powers These Products?



NEURAL NETWORKS

power many features in these products

TODAY: 24 Nov 2026

A day before the Final Test.

TODAY: 24 Nov 2026

A day before the Final Test.

NOVEMBER 2026

SUN	MON	TUE	WED	THU	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25 FINAL TEST	26	27	28
29	30					

TODAY: 24 Nov 2026

A day before the Final Test.

NOVEMBER 2026

SUN	MON	TUE	WED	THU	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

FINAL TEST

You are worried about your test tomorrow

TODAY: 24 Nov 2026

A day before the Final Test.

NOVEMBER 2026

SUN	MON	TUE	WED	THU	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25 FINAL TEST	26	27	28
29	30					

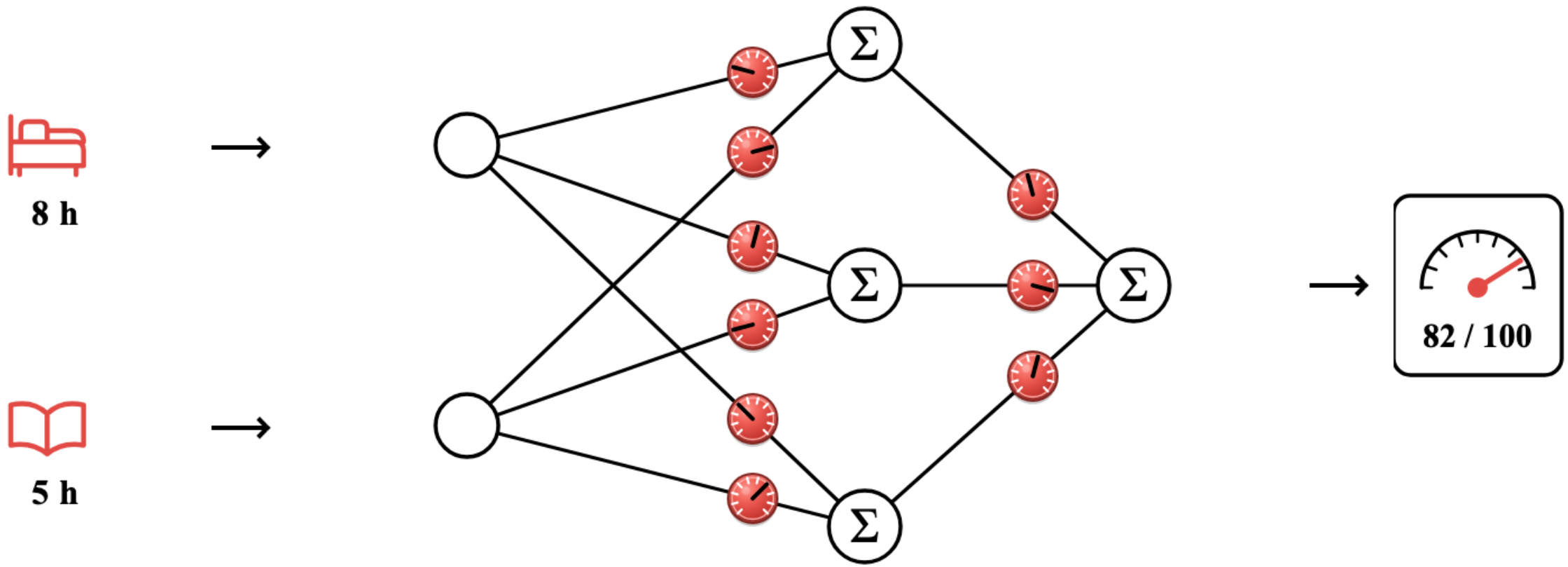
You are worried about your test tomorrow

Someone Gives You This Device

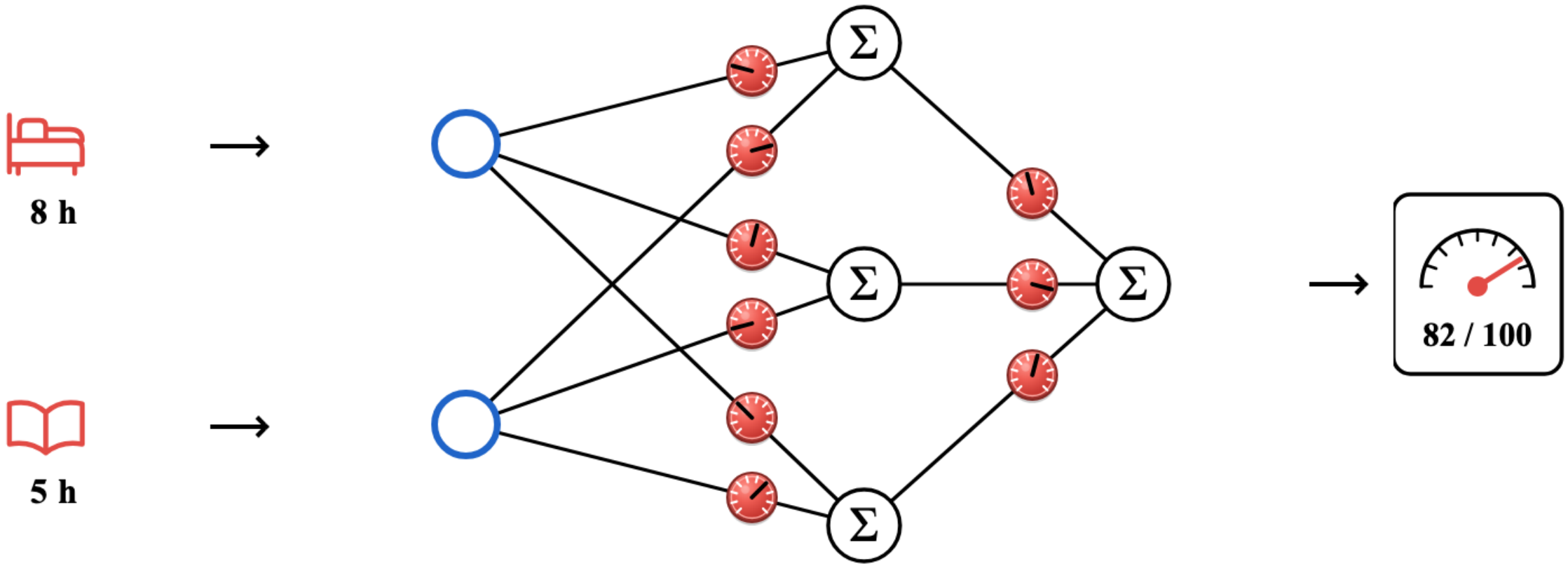
Someone Gives You This Device

You don’t know how it works.
Would you trust it?

Neural Network Inside!

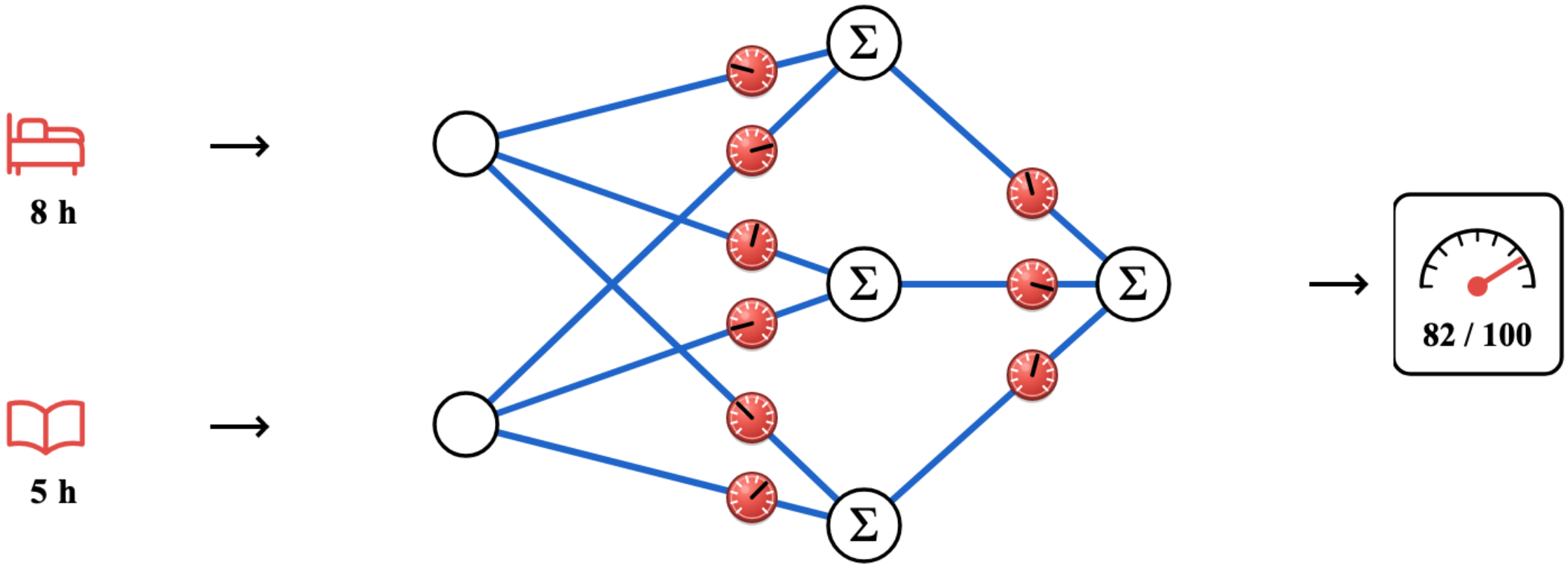


Neural Network Inside!



1 **INPUT NODES**
receive the input values

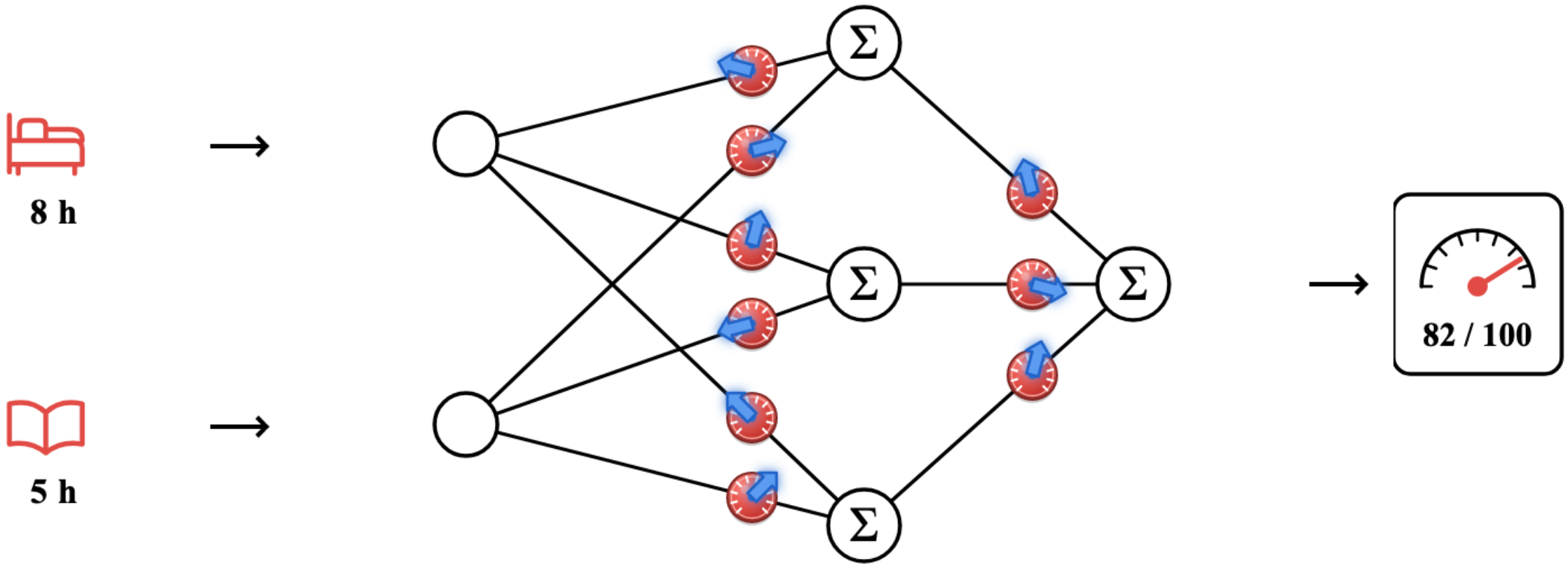
Neural Network Inside!



1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

Neural Network Inside!



1 INPUT NODES
receive the input values

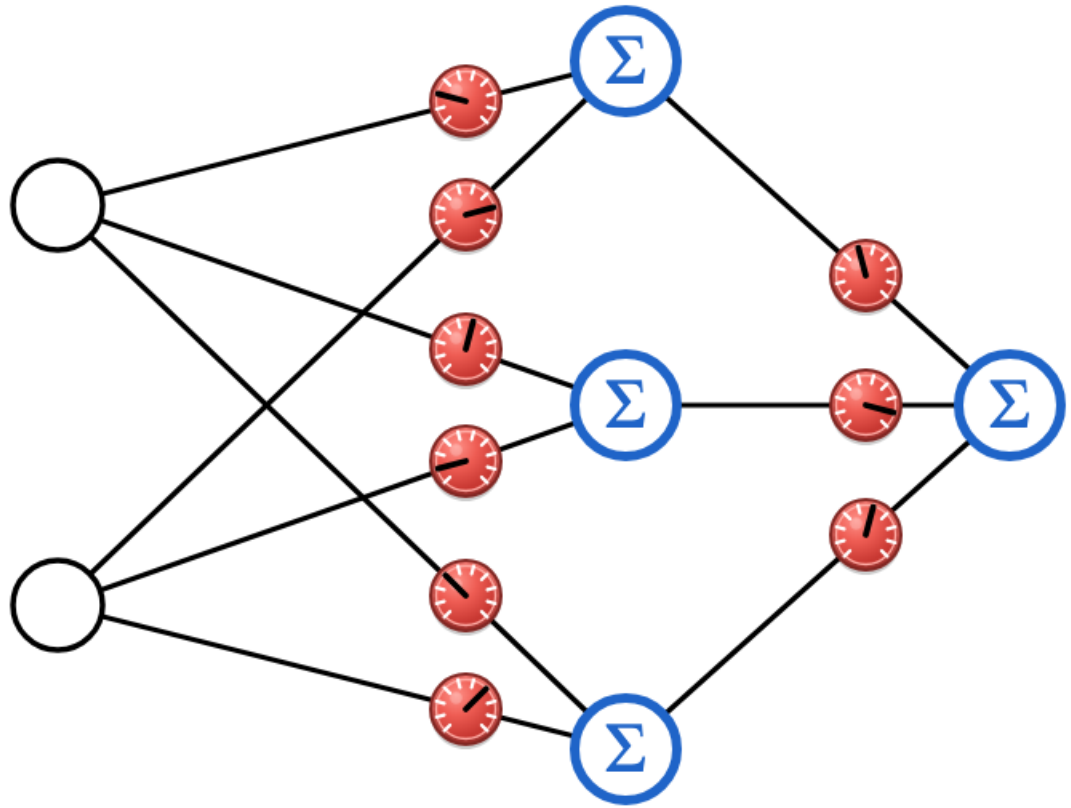
2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

Neural Network Inside!

8 h

5 h



82 / 100

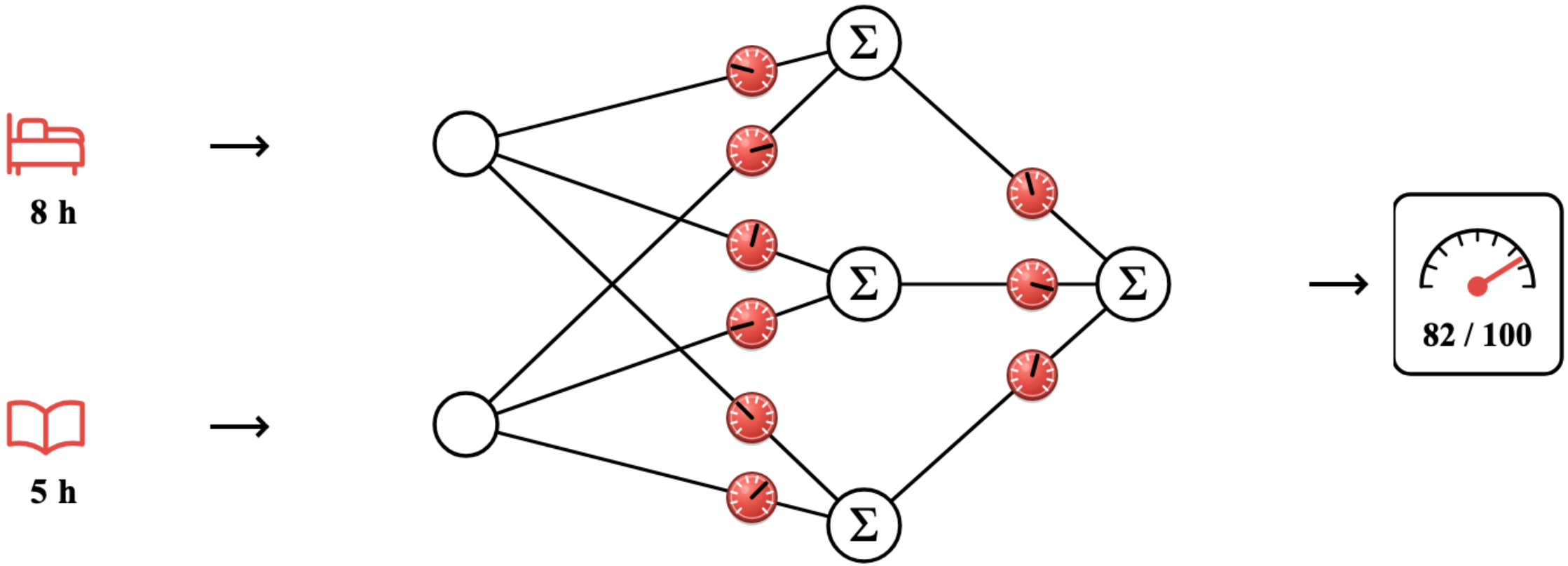
1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

4 ARTIFICIAL NEURONS
combine inputs and produce one output

Neural Network Inside!



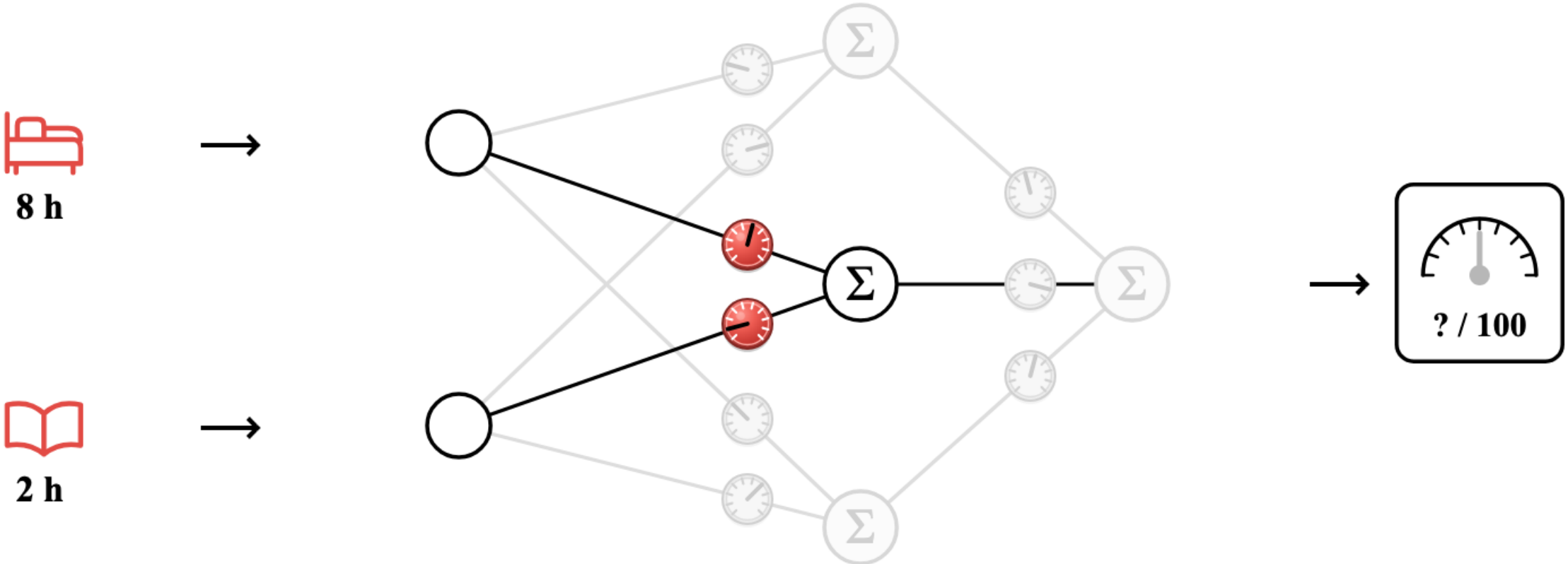
1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

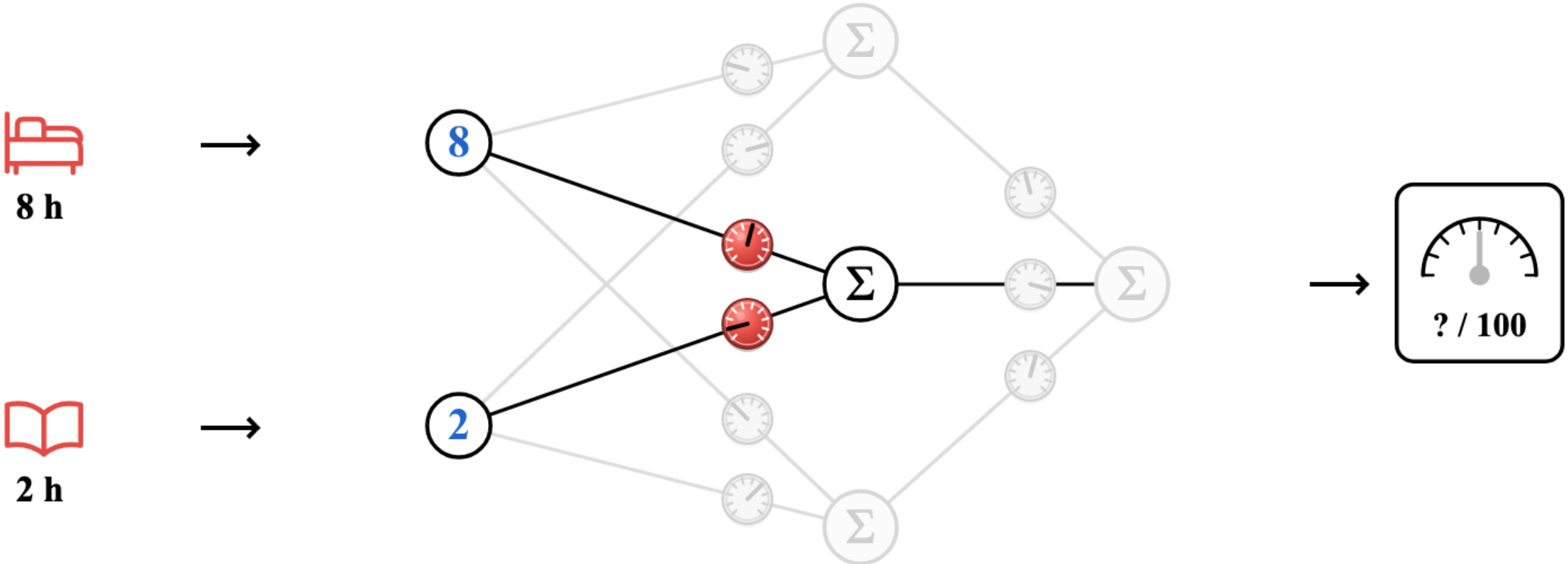
3 WEIGHTS
strengthen or weaken each signal

4 ARTIFICIAL NEURONS
combine inputs and produce one output

Inference: Use the Network

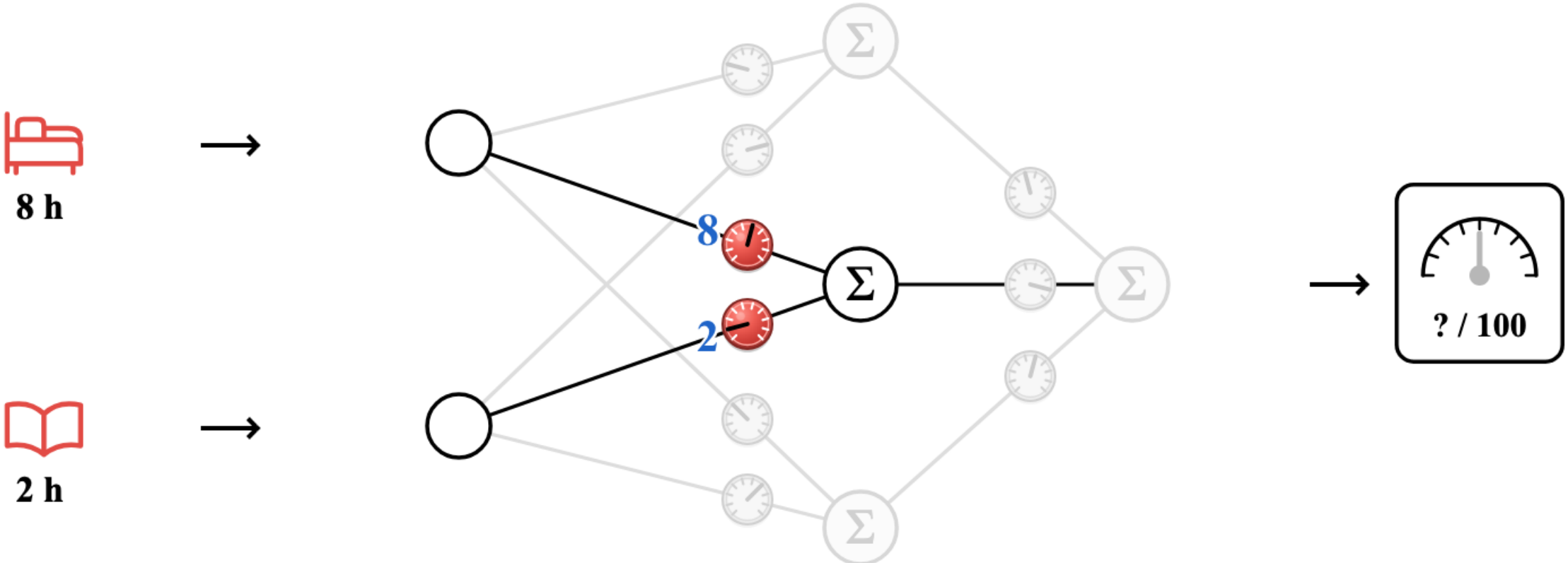


Inference: Use the Network



1 **INPUT NODES**
receive the input values

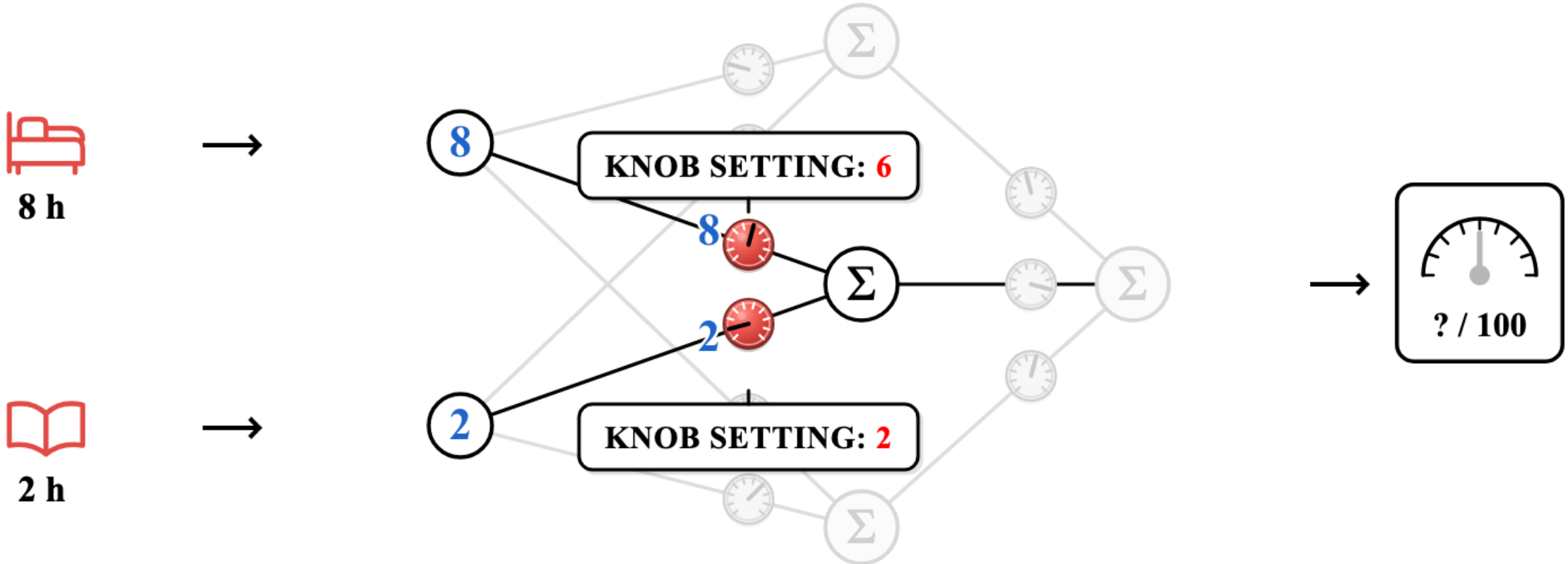
Inference: Use the Network



1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

Inference: Use the Network

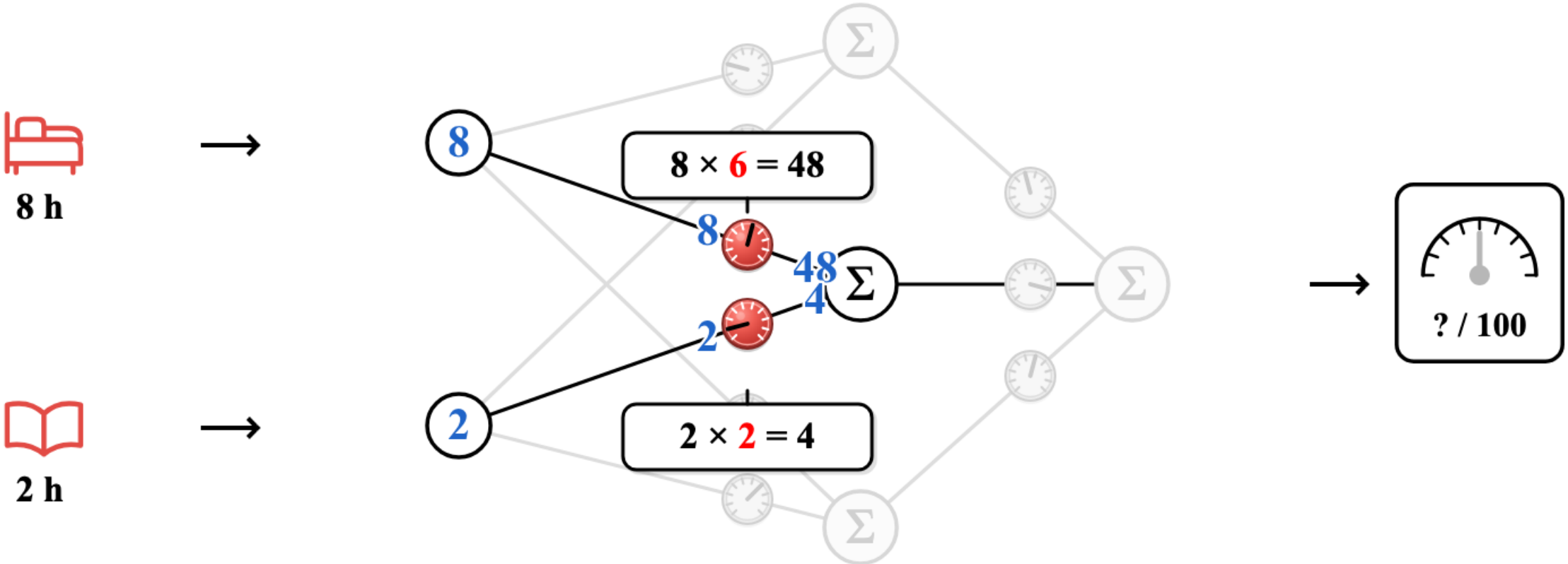


1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

Inference: Use the Network

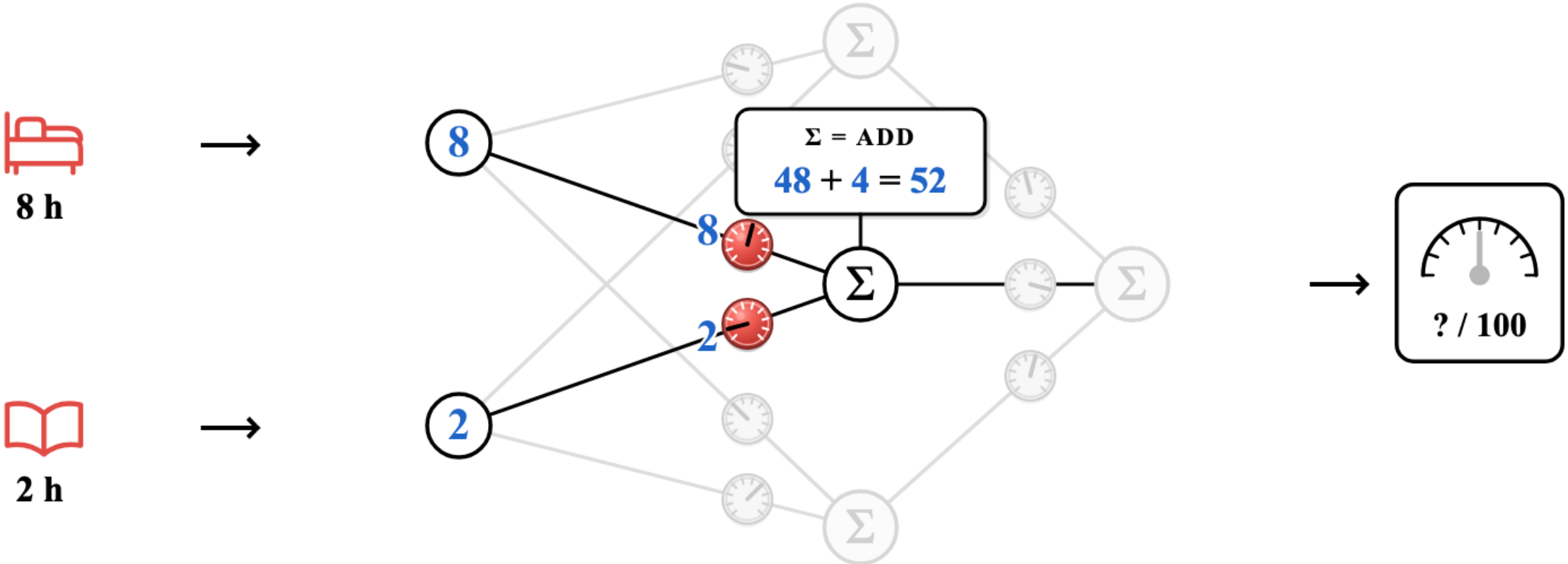


1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

Inference: Use the Network



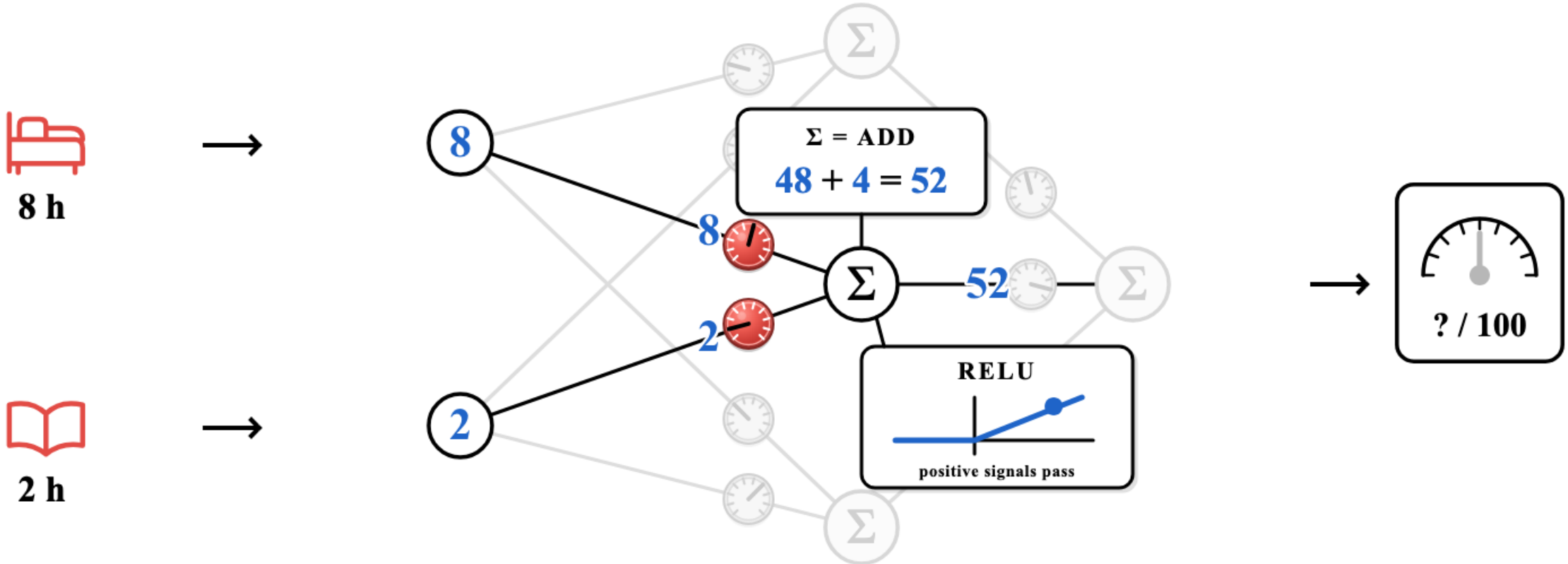
1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

4 ARTIFICIAL NEURONS
combine inputs and produce one output

Inference: Use the Network



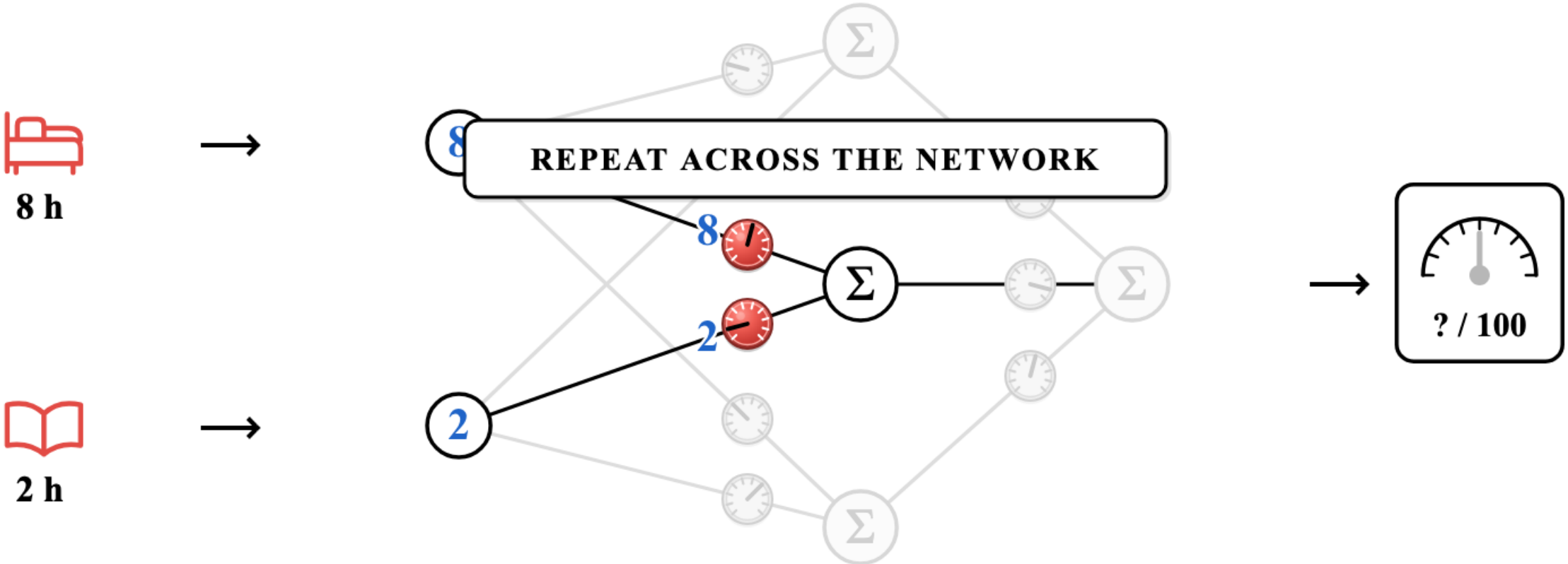
1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

3 WEIGHTS
strengthen or weaken each signal

4 ARTIFICIAL NEURONS
combine inputs and produce one output

Inference: Use the Network



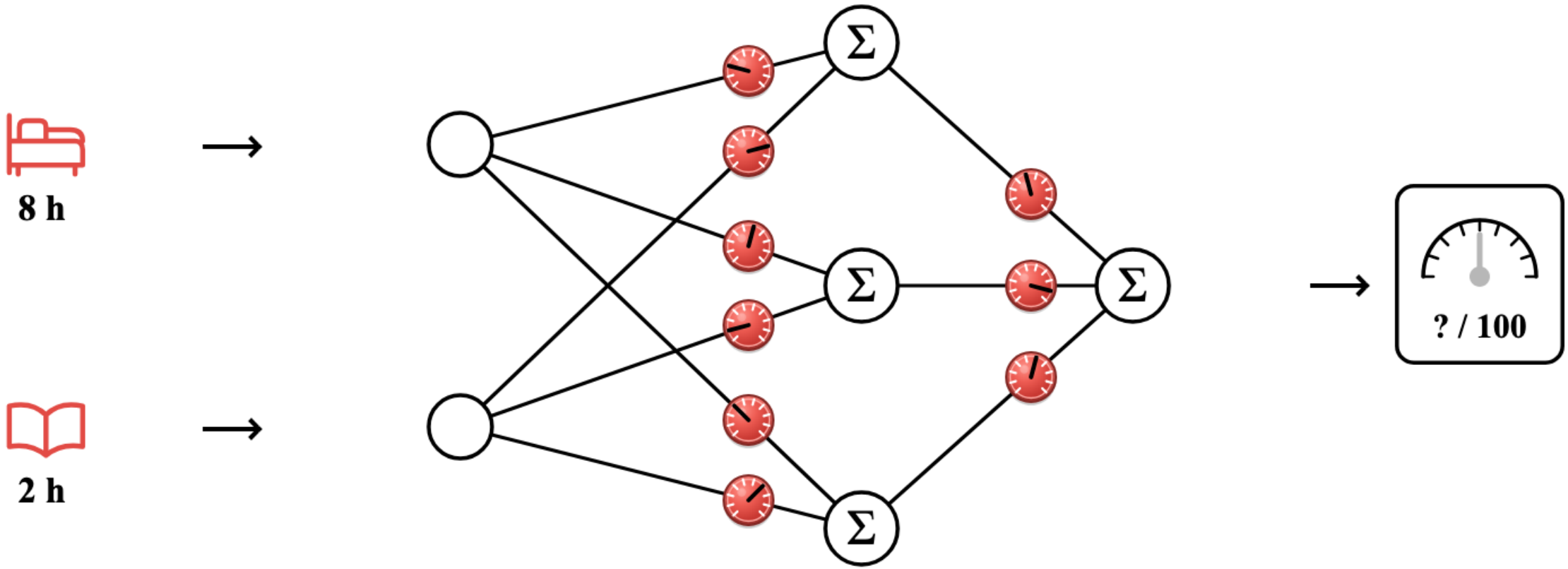
1 INPUT NODES
receive the input values

2 SYNAPSES
carry signals between nodes

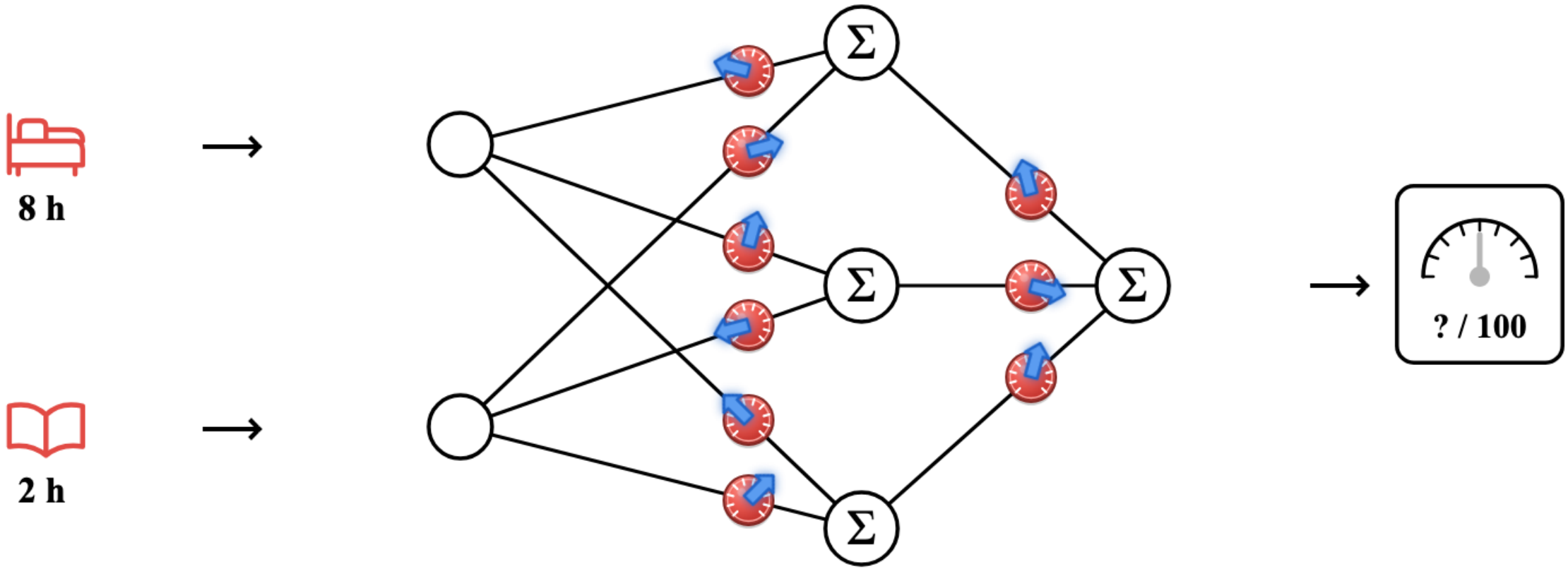
3 WEIGHTS
strengthen or weaken each signal

4 ARTIFICIAL NEURONS
combine inputs and produce one output

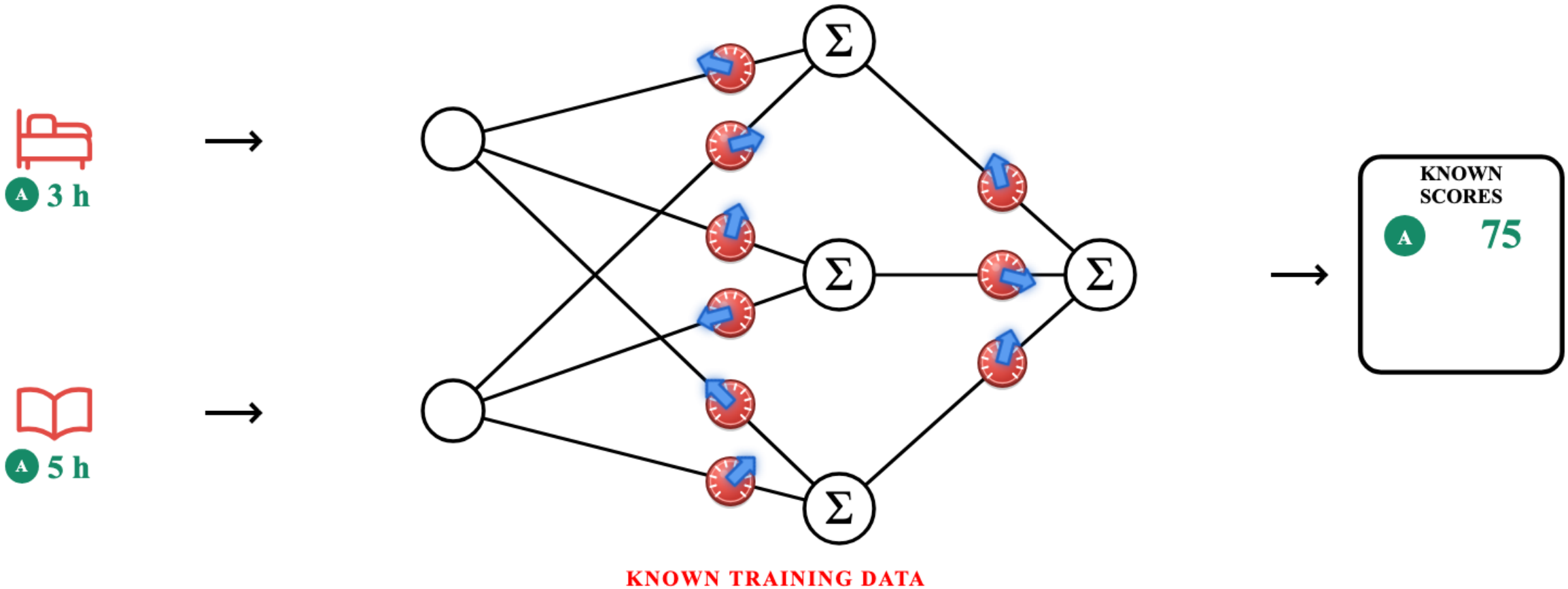
So Far, the Knob Positions Were Already Set



Training Determines the Knob Positions



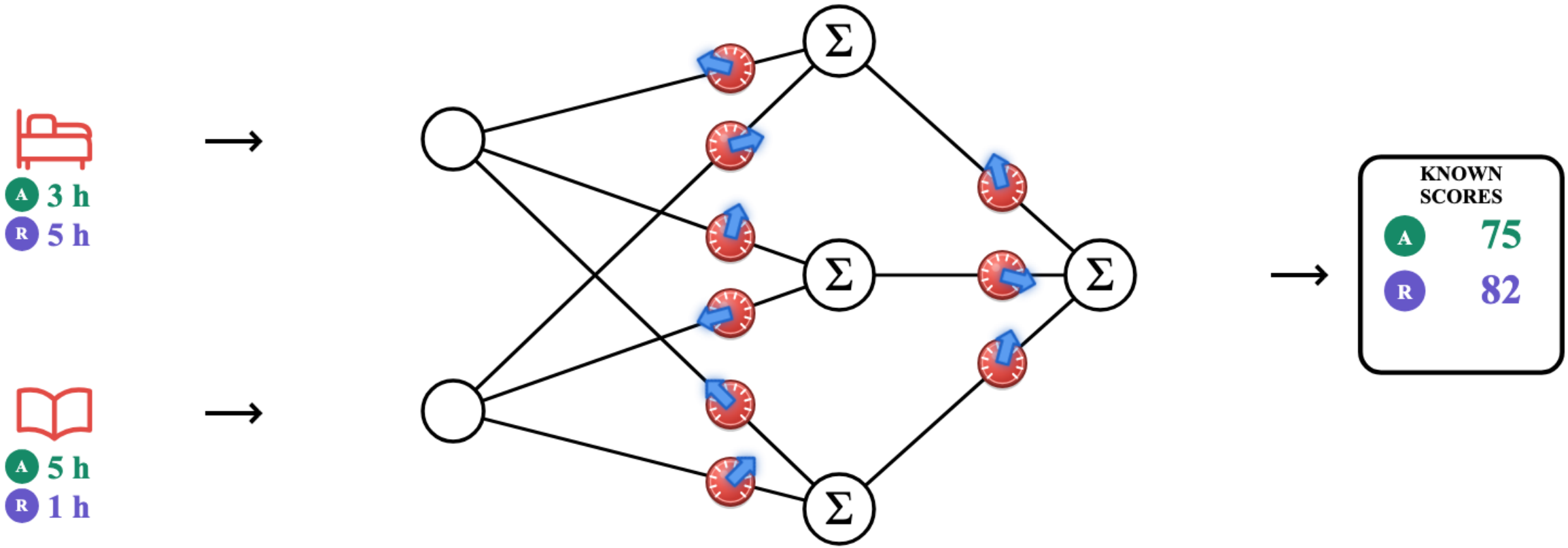
Training Determines the Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75

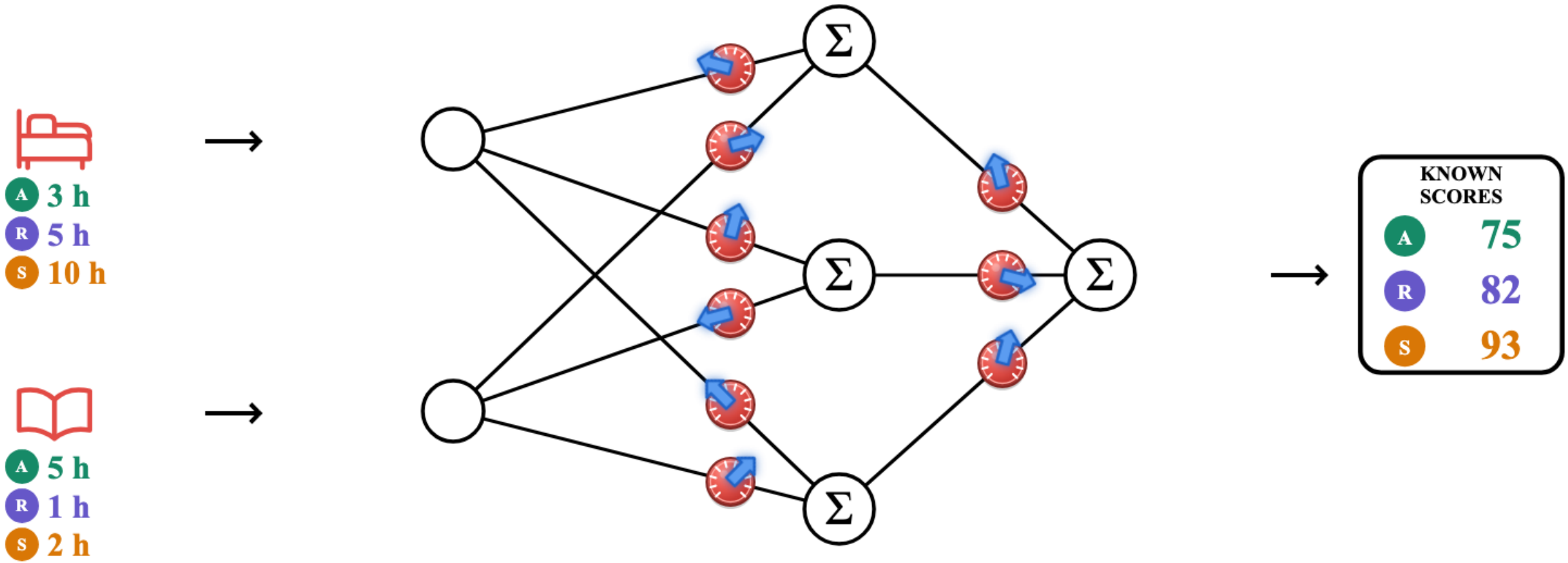
Training Determines the Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82

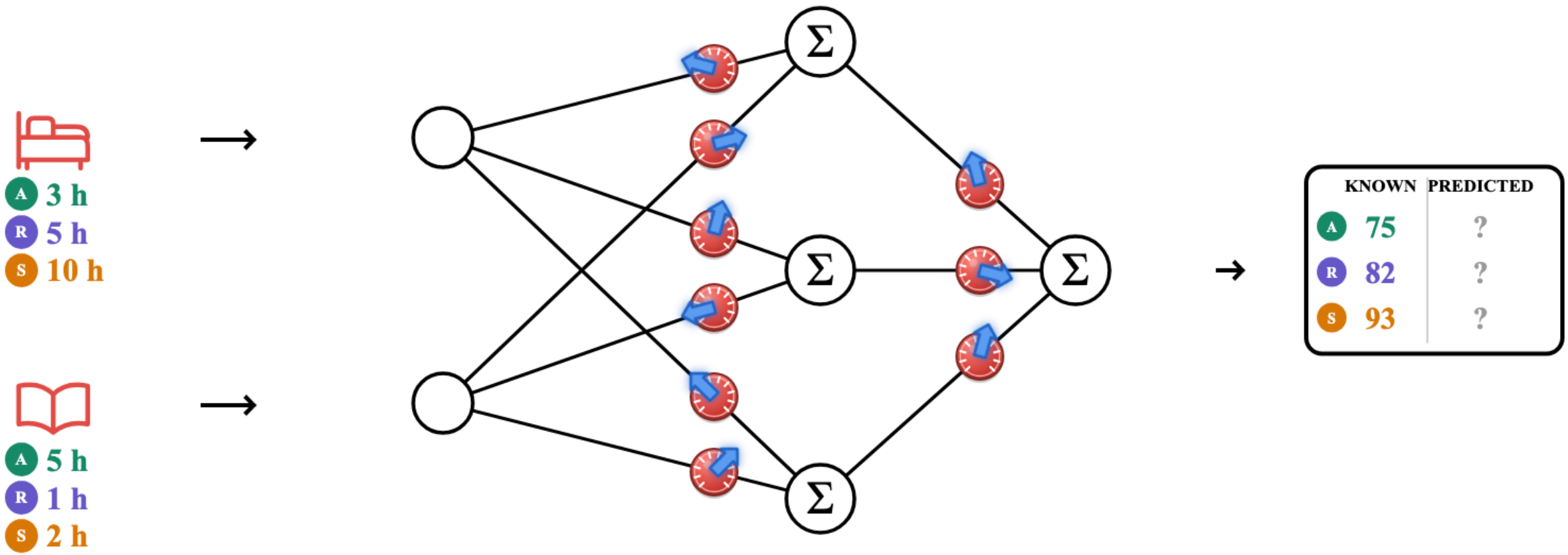
Training Determines the Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82
S	**Sunny Choi** 10h sleep · 2h study	SCORE 93

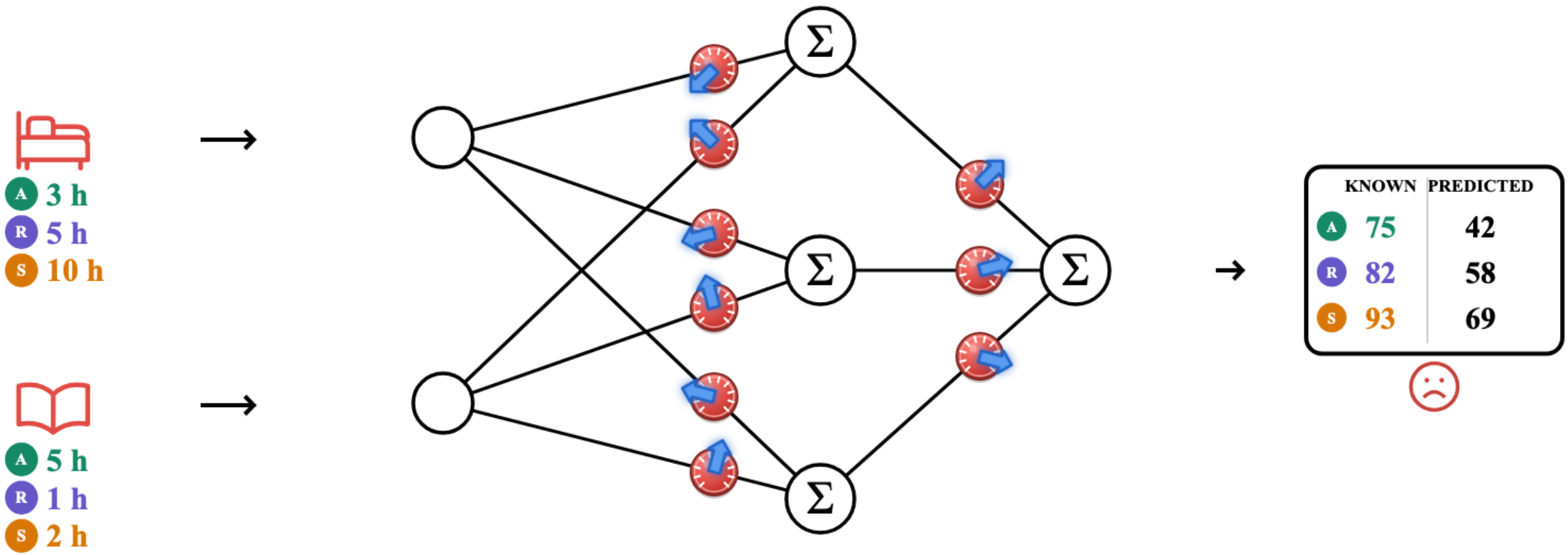
Training: Search for Better Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82
S	**Sunny Choi** 10h sleep · 2h study	SCORE 93

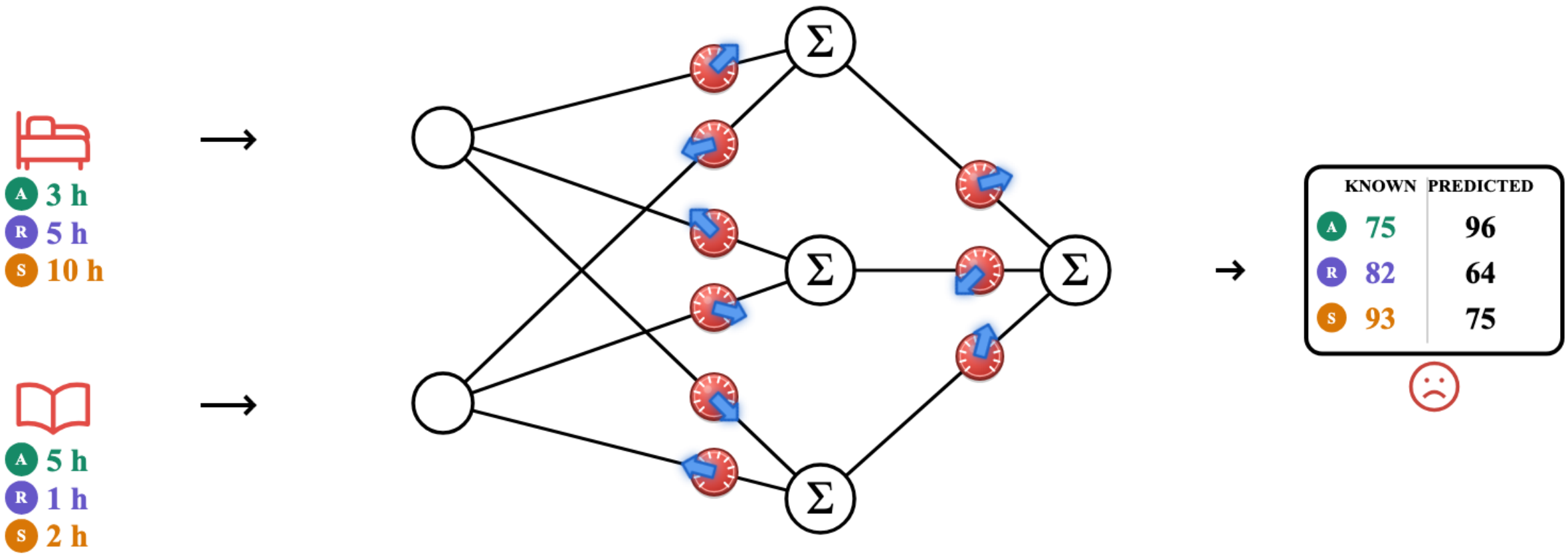
Training: Search for Better Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82
S	**Sunny Choi** 10h sleep · 2h study	SCORE 93

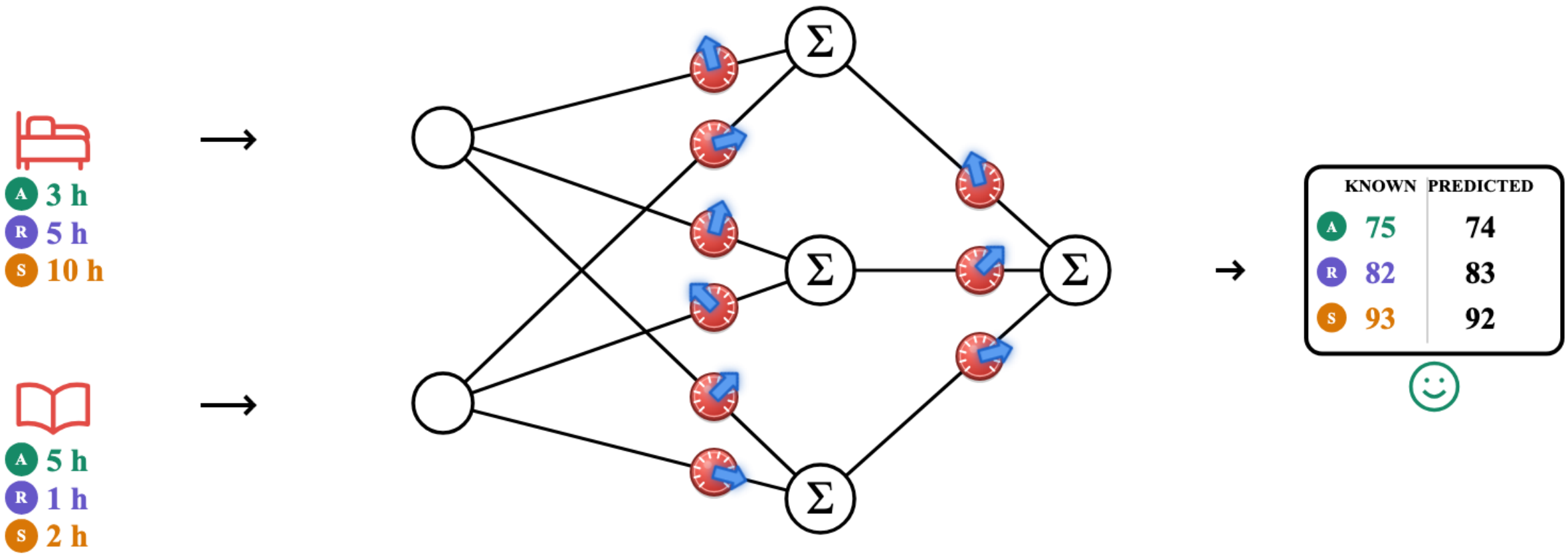
Training: Search for Better Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82
S	**Sunny Choi** 10h sleep · 2h study	SCORE 93

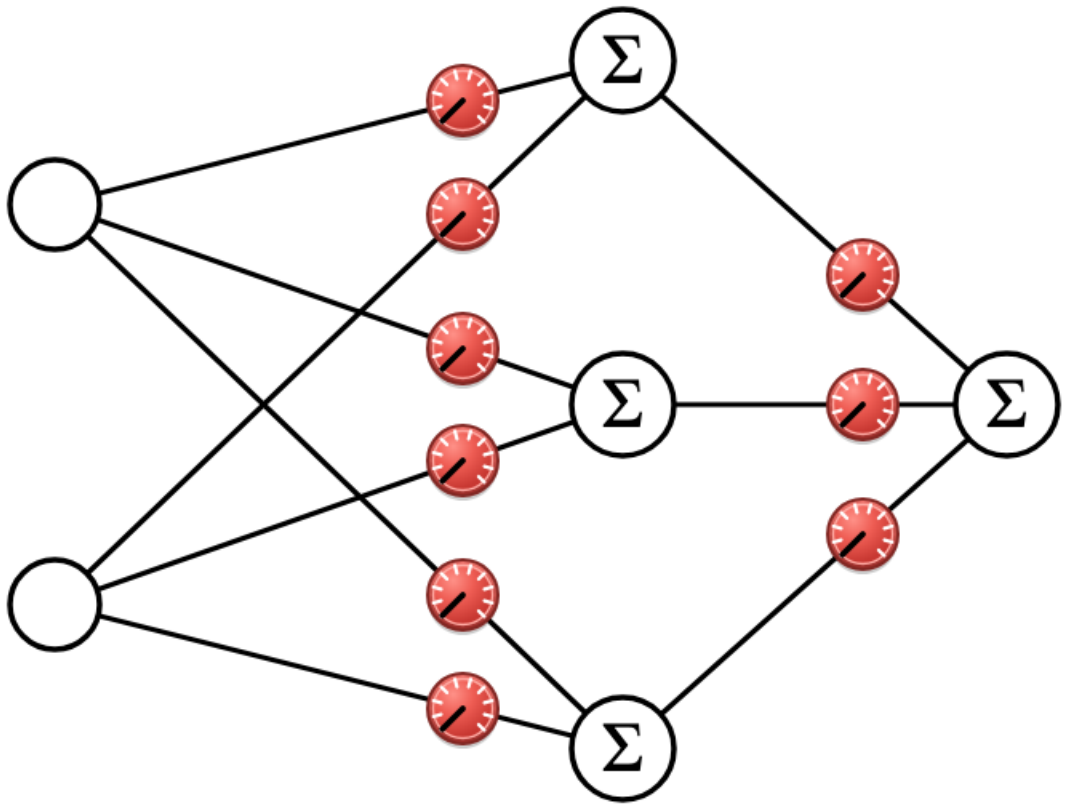
Training: Search for Better Knob Positions



KNOWN TRAINING DATA

A	**Apple Cheung** 3h sleep · 5h study	SCORE 75
R	**Rainbow Wong** 5h sleep · 1h study	SCORE 82
S	**Sunny Choi** 10h sleep · 2h study	SCORE 93

How Can We Search?



How Can We Search?

9 KNOBS · 10 POSITIONS EACH

How Can We Search?

9 KNOBS · 10 POSITIONS EACH

TRY

How Can We Search?

9 KNOBS · 10 POSITIONS EACH

TRY

---

10 TRIES / SECOND → **≈ 3 YEARS · 2 MONTHS · 2 DAYS**

How Can We Search?

9 KNOBS · 10 POSITIONS EACH

TRY

---

10 TRIES / SECOND → **≈ 3 YEARS · 2 MONTHS · 2 DAYS**

OUR NETWORK

9

KNOBS

How Can We Search?

9 KNOBS · 10 POSITIONS EACH

TRY

---

10 TRIES / SECOND → **≈ 3 YEARS · 2 MONTHS · 2 DAYS**

OUR NETWORK

9

KNOBS

VS

DEEPSEEK-V4 PRO

1,700,000,000,000

KNOBS

How Did They Train It?

How Can We Search?

9 KNOBS · 10 POSITIONS EACH

TRY

---

10 TRIES / SECOND → **≈ 3 YEARS · 2 MONTHS · 2 DAYS**

OUR NETWORK

9

KNOBS

VS

DEEPSEEK-V4 PRO

1,700,000,000,000

KNOBS

How Did They Train It?

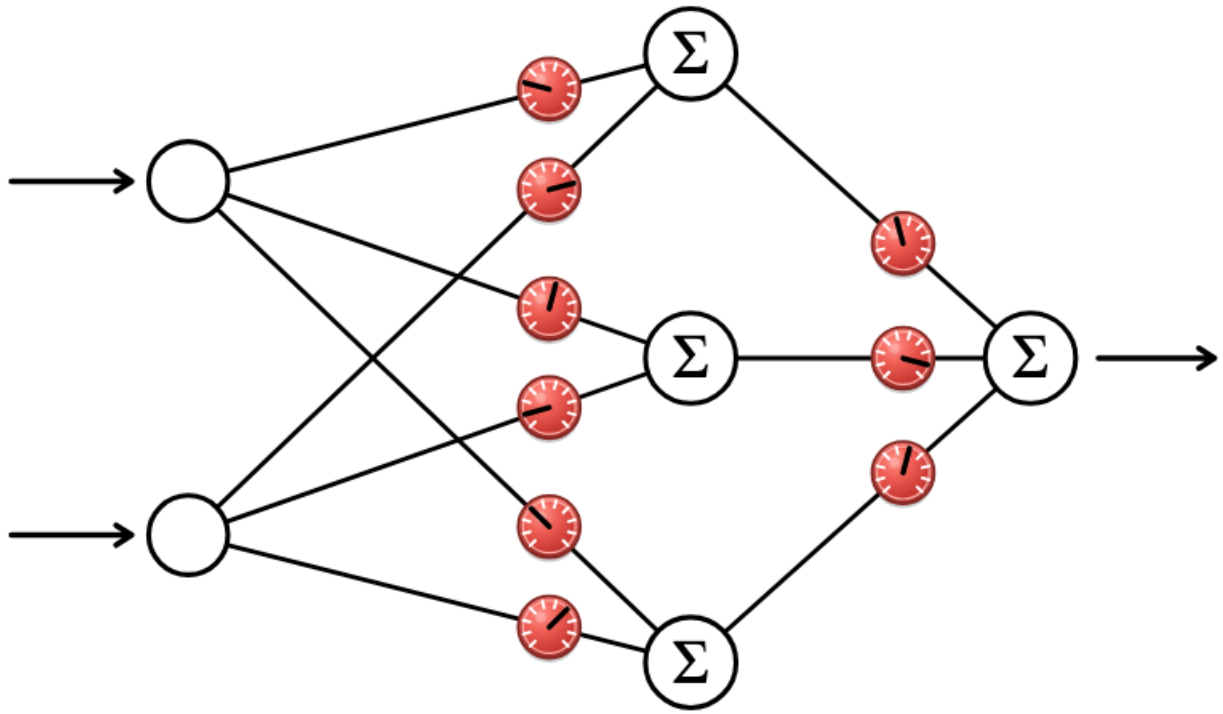
BACKPROPAGATION

provides a useful direction for updating all the knobs AT THE SAME TIME

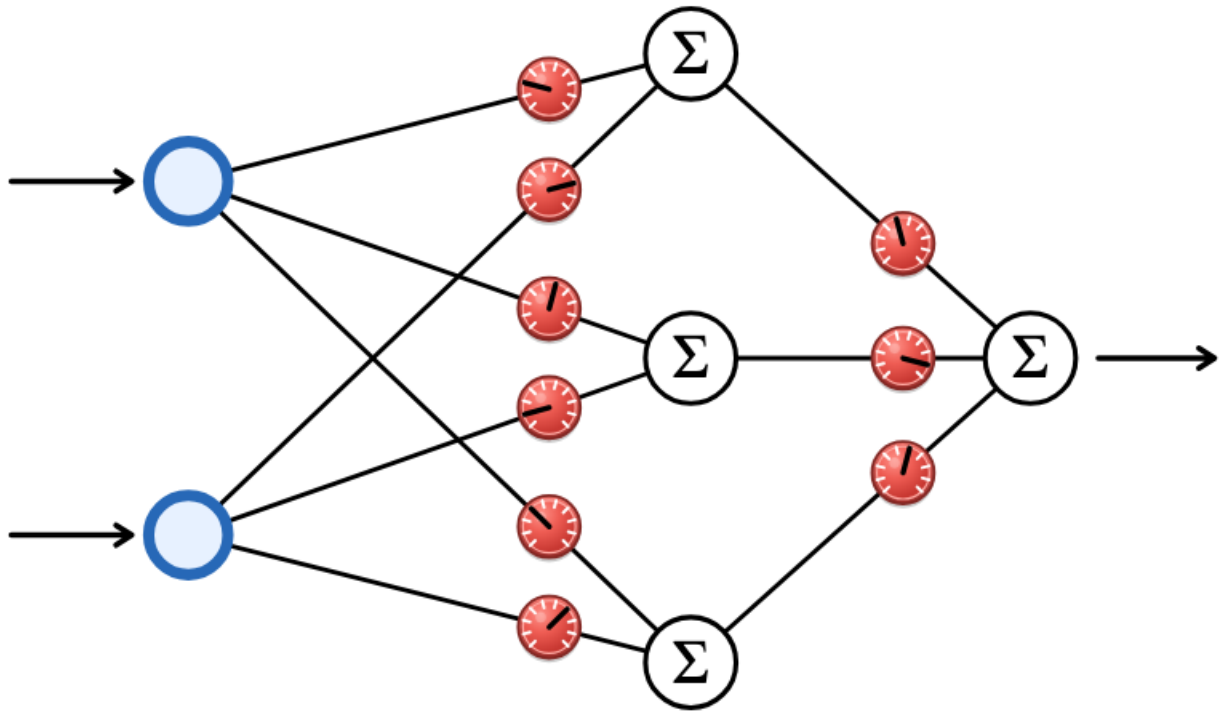
See how training data changes the knob positions

DEMO 3 · Training in Action

Layers of a Neural Network



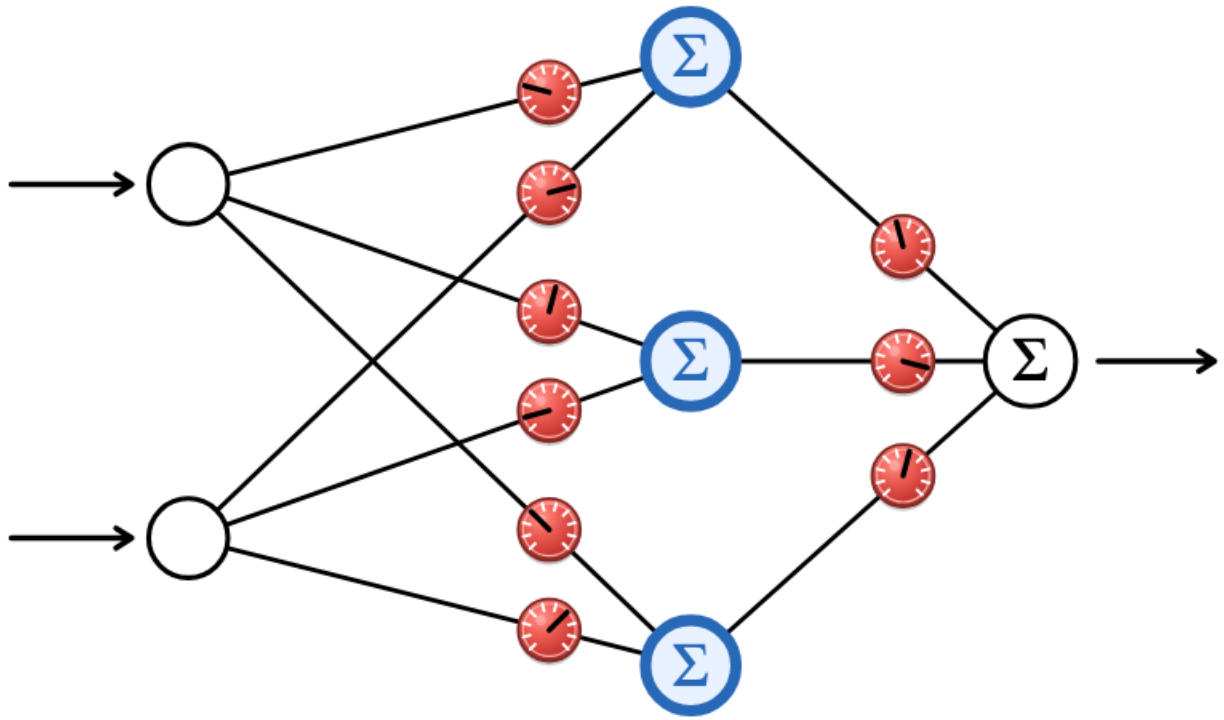
Layers of a Neural Network



1 **INPUT LAYER**

2 input nodes

Layers of a Neural Network



1 **INPUT LAYER**

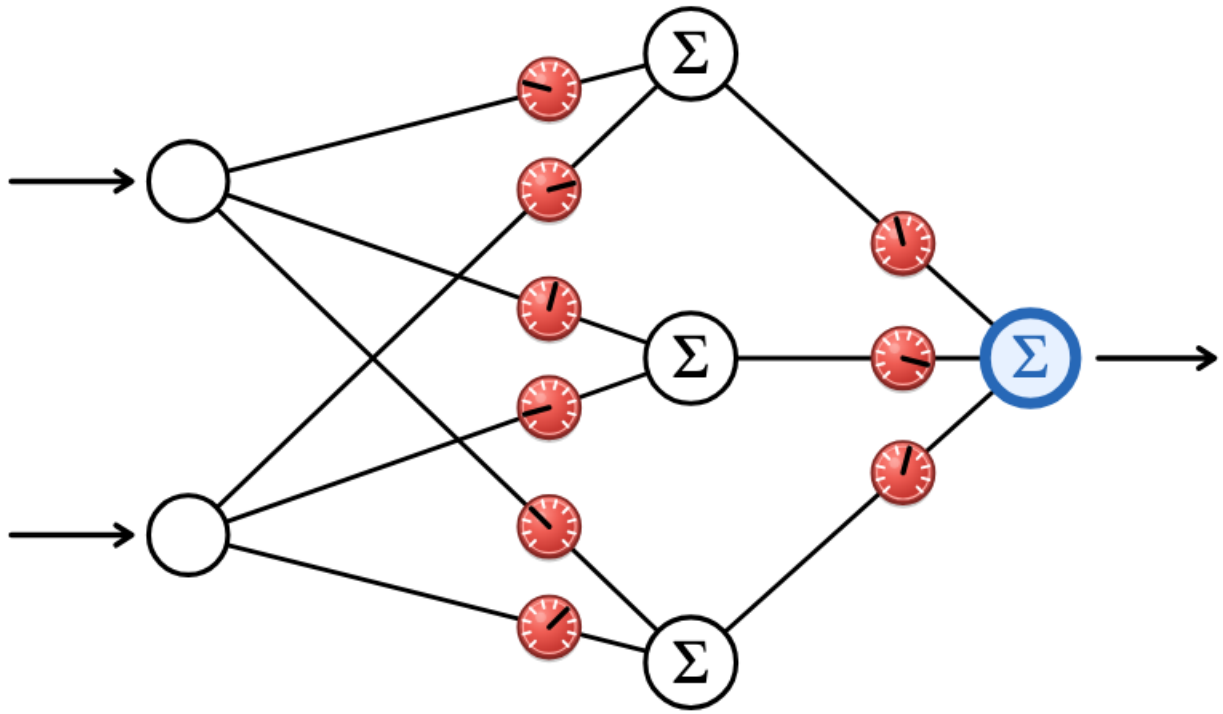
2 input nodes

2 **HIDDEN LAYER(S)**

1 hidden layer · 3 neurons

2+ hidden layers → deep network

Layers of a Neural Network

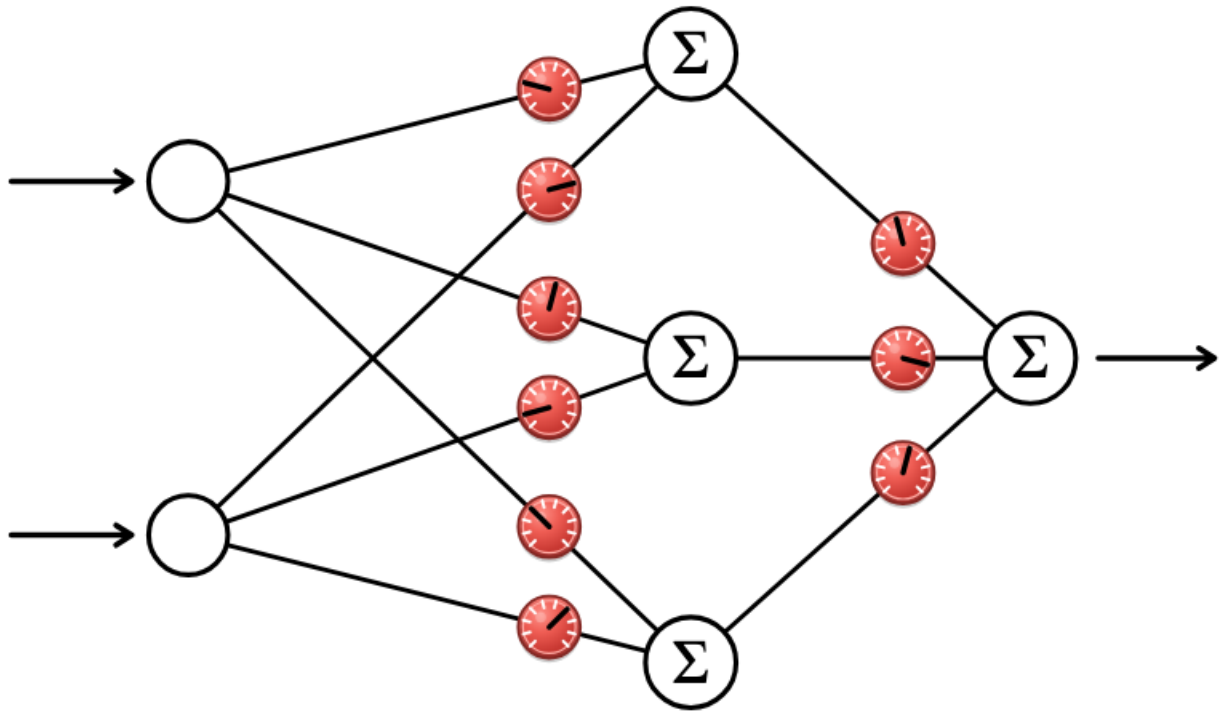


1 **INPUT LAYER**
2 input nodes

2 **HIDDEN LAYER(S)**
1 hidden layer · 3 neurons

3 **OUTPUT LAYER**
1 output neuron

Layers of a Neural Network

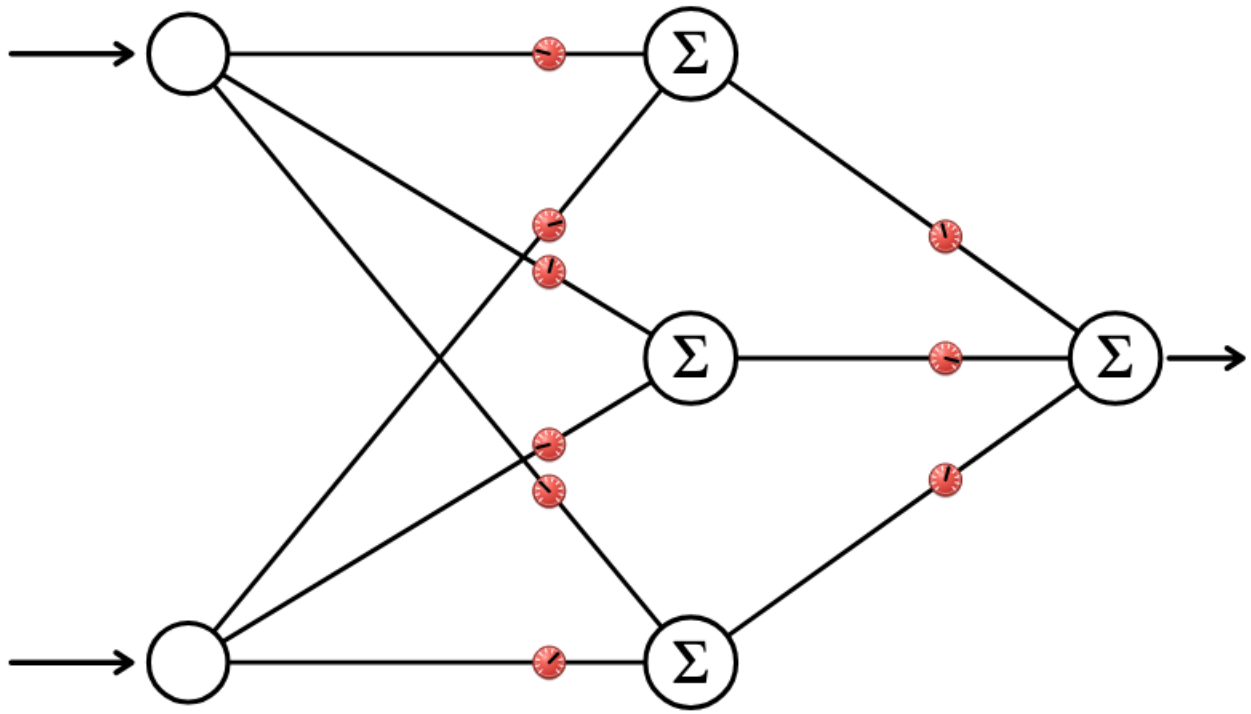


1 **INPUT LAYER**
2 input nodes

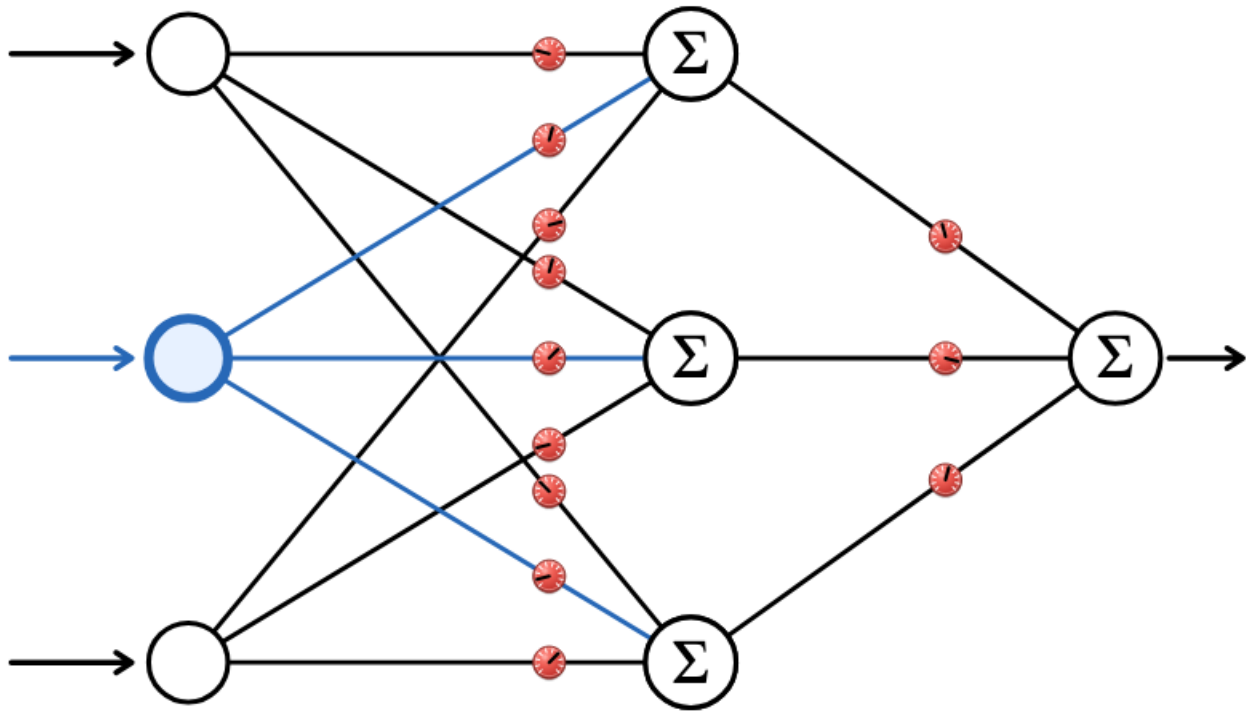
2 **HIDDEN LAYER(S)**
1 hidden layer · 3 neurons

3 **OUTPUT LAYER**
1 output neuron

Changing the Architecture



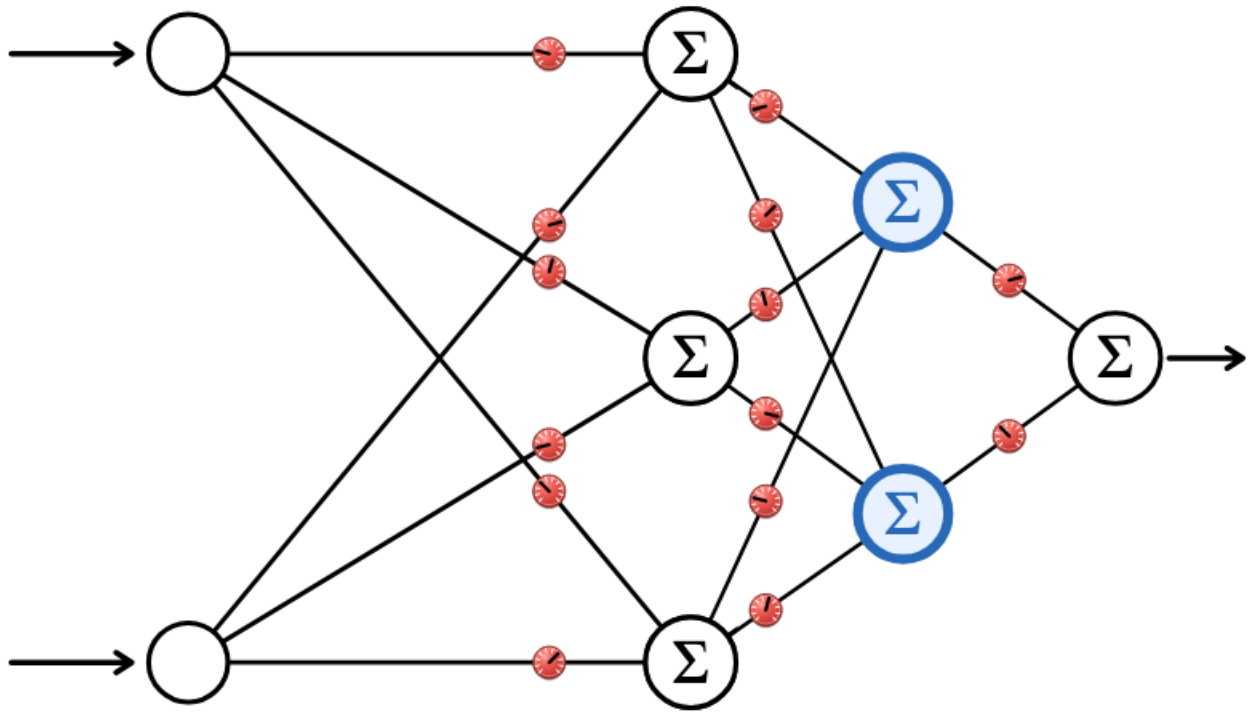
Changing the Architecture



1 **INPUT LAYER**

2 → 3 input nodes

Changing the Architecture



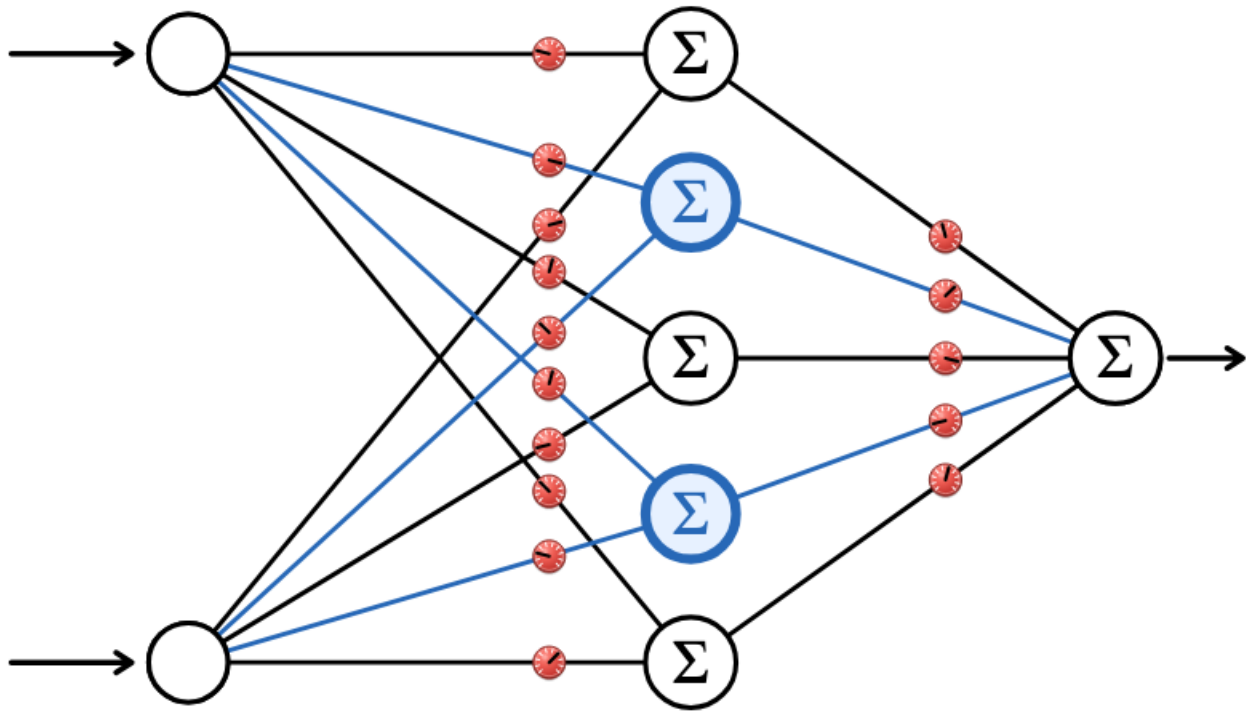
1 **INPUT LAYER**

2 → 3 input nodes

2 **HIDDEN LAYER(S)**

1 → 2 hidden layers

Changing the Architecture



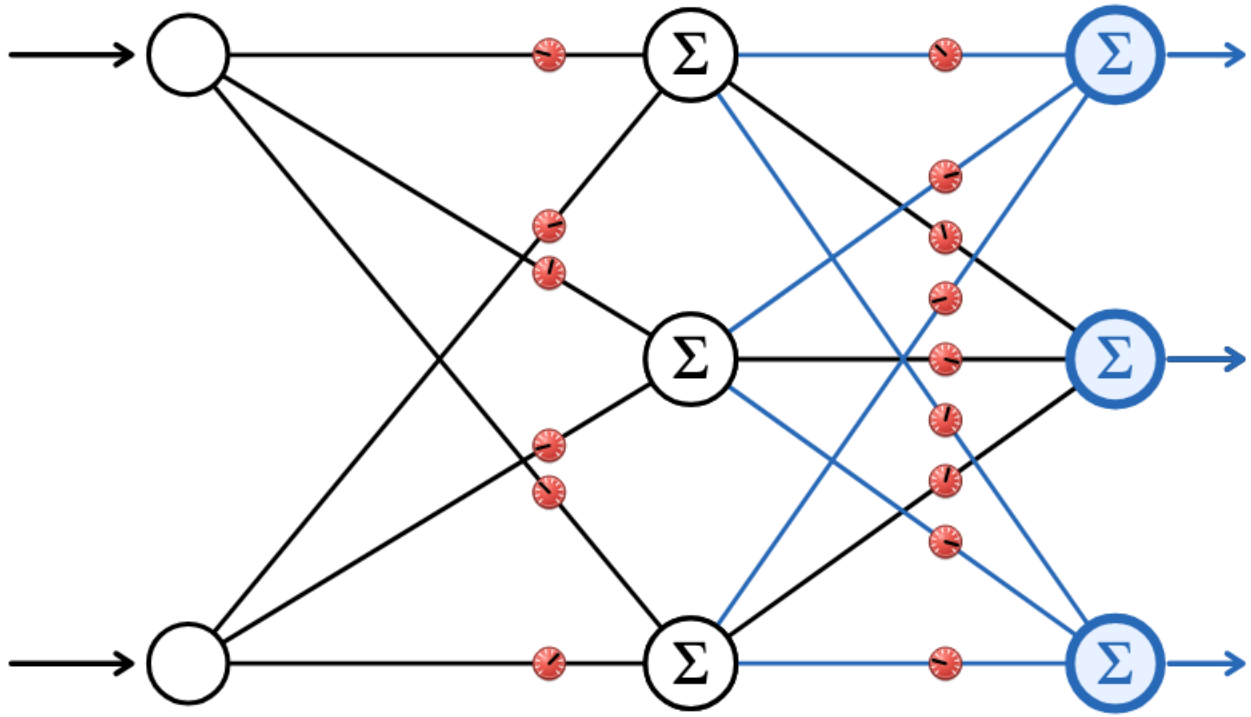
1 **INPUT LAYER**

2 → 3 input nodes

2 **HIDDEN LAYER(S)**

3 → 5 neurons / layer

Changing the Architecture



1 **INPUT LAYER**
2 → 3 input nodes

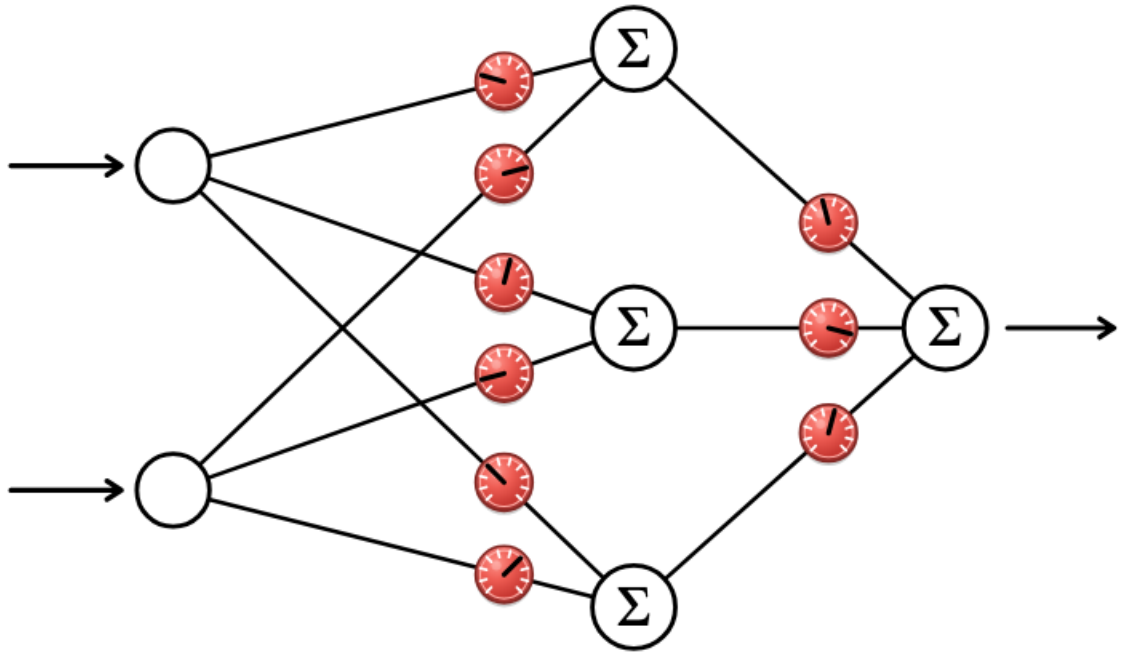
2 **HIDDEN LAYER(S)**
3 → 5 neurons / layer

3 **OUTPUT LAYER**
1 → 3 output neurons

Different Inputs, Different Outputs

NUMBERS IN

SLEEP	8 h
STUDY	3 h



NUMBER OUT

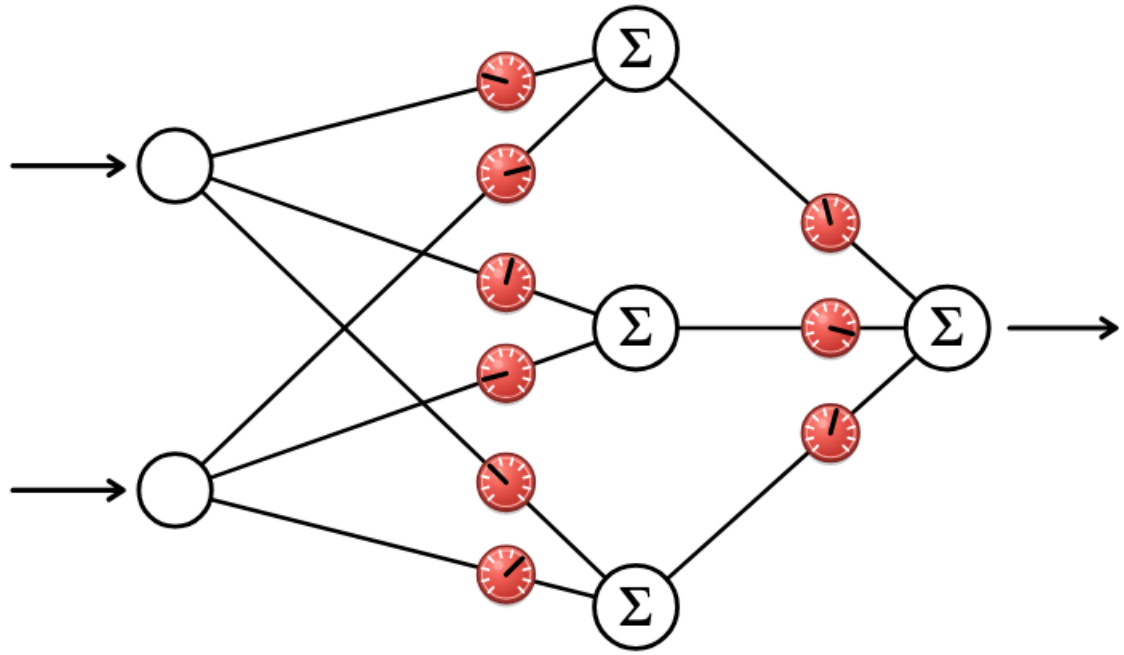
74 / 100

PREDICTED SCORE

OUR EXAMPLE numbers → predicted score

Different Inputs, Different Outputs

LOW-RES IMAGE IN



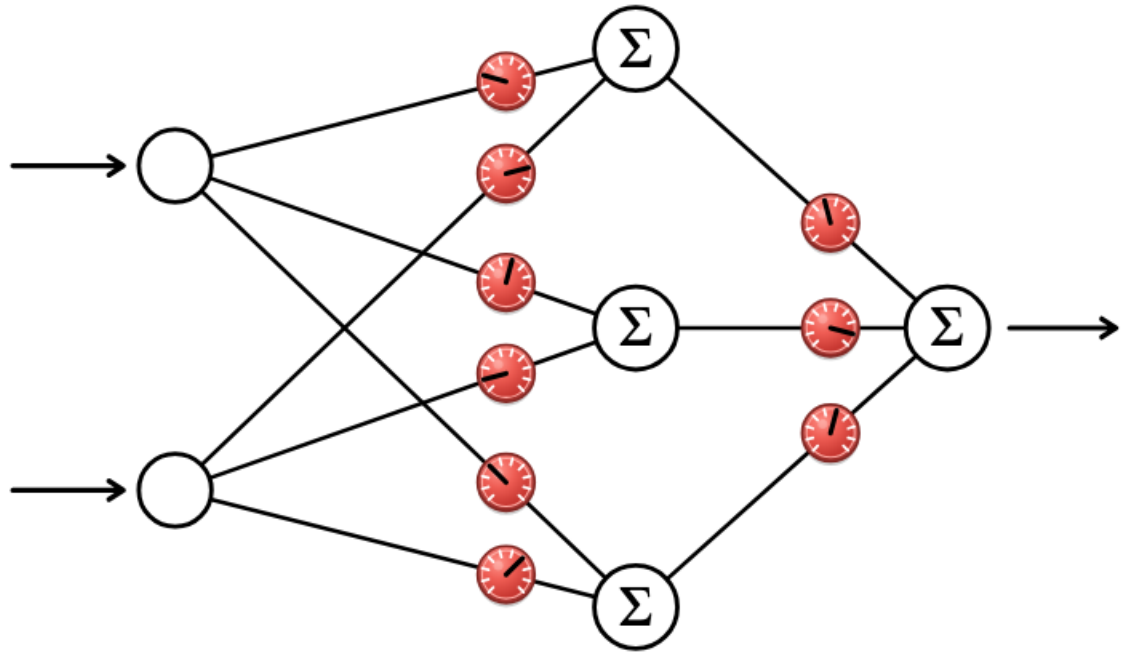
ENHANCED IMAGE OUT

SUPER-RESOLUTION low-resolution image → high-resolution image

Different Inputs, Different Outputs

WORDS IN

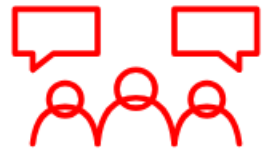
The deadline was



NEXT WORD OUT

yesterday

TEXT PREDICTION previous words → next word



1. **Ed Login**
   on your phone
2. **Camera App**
3. **Scan QR**
4. **Submit**

ACTIVITY · EXIT TICKET