

Demystifying Artificial Intelligence Sorcery

(Part 1: Fuzzy Logic & Neural Networks)^a

Abdelbacet Mhamdi abdelbacet.mhamdi@bizerte.r-iset.tn

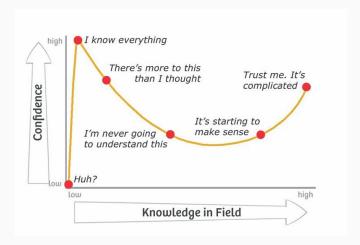
Dr.-Ing. in Electrical Engineering Senior Lecturer at ISET Bizerte

[&]quot;Available @ https://github.com/a-mhamdi/jlai/



This document features some materials gathered from multiple online sources. Please note no copyright infringement is intended, and I do not own nor claim to own any of the original materials. They are used for educational purposes only. I have included links solely as a convenience to the reader. Some links within these slides may lead to other websites, including those operated and maintained by third parties. The presence of such a link does not imply a responsibility for the linked site or an endorsement of the linked site, its operator, or its contents.

DUNNING-KRUGER EFFECT

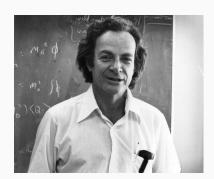


Kruger, J. and Dunning, D. (1999) *Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments.* **J Pers Soc Psychol.** 77(6) pp. 1121–1134.

2/79 A. MHAMDI Demystifying Al Sorcery

"Knowledge isn't free. You have to pay attention."

Richard P. Feynman

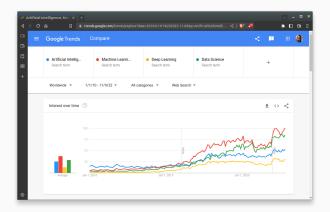


ROADMAP

- 1. An overview
- 2. Fuzzy Logic
- 3. Neural Networks
- 4. Quizzes

An overview

TRENDS



"Numbers represent search interest relative to the highest point on the chart for the given region and time.

- A value of 100 is the peak popularity for the term;
- A value of 50 means that the term is half as popular;
- A score of 0 means there was not enough data for this term."

TOP USES

























Artificial intelligence is a branch of computer science which focuses on automation of intelligent behavior.

SOME DEFINITIONS CAN BE CATEGORIZED INTO FOUR FRAMES.

Artificial intelligence is a branch of computer science which focuses on automation of intelligent behavior.



SYSTEMS THAT THINK LIKE HUMANS

[Bel78]

"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem-solving, learning..."

Bellman, R. E. An Introduction to Artificial Intelligence: Can Computers Think? Boyd & Fraser Publishing Company.

[Hau89]

"The exciting new effort to make computers think[...] machines with minds, in the full and literal sense"

Haugeland, J. (1989). Artificial Intelligence: The Very Idea. A Bradford book. MIT Press.

Demystifying Al Sorcery 8/79 Δ ΜΗΔΜΟΙ

[CMM85]

"The study of mental faculties through the use of computational models."

Charniak, E., McDermott, D., and McDermott, D. V. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley series in computer science and information processing. Addison-Wesley.

[Win92]

"The study of the computations that make it possible to perceive, reason, and act."

Winston, P. H. (1992). Artificial Intelligence. A-W Series in Computer Science. Addison-Wesley Publishing Company.

9/79

SYSTEMS THAT ACT LIKE HUMANS

[Kur92]

"The art of creating machines that perform functions that require intelligence when performed by people."

Kurzweil, R. (1992). The Age of Intelligent Machines. Viking.

[RK91]

"The study of how to make computers do things at which, at the moment, people are better."

Rich, E. and Knight, K. (1991). Artificial Intelligence. Artificial Intelligence Series. McGraw-Hill.

10/79 A. MHAMDI Demystifying Al Sorcery

Systems that act rationally

[Sch90]

"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes."

Schalkoff, R. J. (1990). Artificial Intelligence: An Engineering Approach. McGraw-Hill Computer science series. McGraw-Hill.

"The branch of computer science that is concerned with the automation of intelligent behavior"

Luger, G. F. and Stubblefield, W. A. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Artificial intelligence. Benjamin/Cummings Publishing Company.

Demystifying Al Sorcery 11/79 Δ ΜΗΔΜΟΙ

THOUGHT-PROVOKING QUESTIONS



How to achieve intelligence on a computer system

THOUGHT-PROVOKING QUESTIONS



How to achieve intelligence on a computer system

What do we mean by "Intelligence"?

- Single faculty or gathering of abilities
- Learned or existing
- What happens when we learn
- Are creativity and intuition measurable
- → Does observable behavior infer to intelligence
- → How knowledge is routed in the human brain

TURING TEST

Alan Turing (1950)

The ability to achieve human level performance in all cognitive tasks, sufficient to fool an interrogator.

- ✓ Natural Language Processing (NLP) (Communicate in human language)
- ✓ Knowledge Representation (Store information)
- ✓ Automated Reasoning (Answer questions & draw conclusions)
- ✓ Machine Learning (ML) (Adapt to new circumstances, detect & extrapolate patterns)

13/79

FORMS OF AI

- ★ Expert Systems (Based on knowledge or rule settings)
- ★ Fuzzy Systems (Based on fuzzy set theory)
- * Artificial Neural Networks
- * Genetic Algorithms
- * Belief Networks
- ★ Hybrid Systems (Combine two or more approaches)

PROGRAMMING LANGUAGE

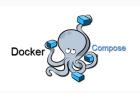




DEVELOPMENT ENVIRONMENTS







- ▲ \$ docker compose up
- ▼ \$ docker compose down







JULIA IN A NUTSHELL

- ▲ **Fast:** native code for multiple platforms via LLVM;
- **Dynamic:** good support for interactive use (like a a scripting language);
- **Reproducible:** environment recreation across platforms, with pre-built binaries;
- **Composable:** multiple dispatch as a paradigm (oop & functional programming);
- General: asynchronous I/O, metaprogramming, debugging, logging; profiling, pkg, ...
- Open Source: GitHub repository at https://github.com/JuliaLang/julia.



JULIA MICRO-BENCHMARKS (1/2)



https://julialang.org/benchmarks



18/79 Demystifying Al Sorcery A. MHAMDI

JULIA MICRO-BENCHMARKS (2/2)

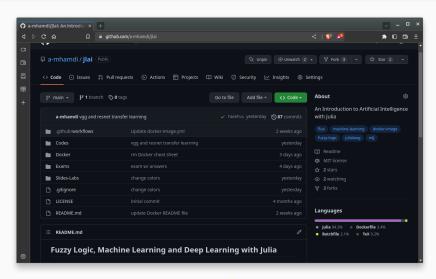
Geometric Means¹ of Micro-Benchmarks by Language

| С | 1.0 |
|-------------|--|
| Julia | 1.17006 |
| LuaJIT | 1.02931 |
| Rust | 1.0999 |
| Go | 1.49917 |
| Fortran | 1.67022 |
| Java | 3.46773 |
| JavaScript | 4.79602 |
| Matlab | 9.57235 |
| Mathematica | 14.6387 |
| Python | 16.9262 |
| R | 48.5796 |
| Octave | 338.704 |
| | Julia Lua]IT Rust Go Fortran Java JavaScript Matlab Mathematica Python R |



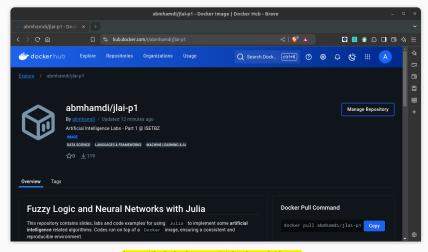
SOURCE CONTROL MANAGEMENT (SCM)





https://github.com/a-mhamdi/jlai

CONTINUOUS INTEGRATION (CI)



https://hub.docker.com/r/abmhamdi/jlai-p1

Fuzzy Logic

WHAT IS FUZZY LOGIC?

"There are many misconceptions about fuzzy logic. To begin with, fuzzy logic is not fuzzy. Basically, fuzzy logic is a precise logic of imprecision. [...] fuzzy logic is designed to deal with imperfect information. Imperfect information is information which in one or more aspects is imprecise, uncertain, incomplete, unreliable, vague or partially true. In the real world, such information is the norm rather than exception."

Lotfi Zadeh, WCECS 2014



"Fuzzy Logic, in computer science, is a form of logic used in some expert systems and other artificial-intelligence applications in which variables can have degrees of truthfulness or falsehood represented by a range of values between 1 (true) and 0 (false). With fuzzy logic, the outcome of an operation can be expressed as a probability rather than as a certainty. For example, in addition to being either true or false, an outcome might have such meanings as probably true, possibly true, possibly false, and probably false."

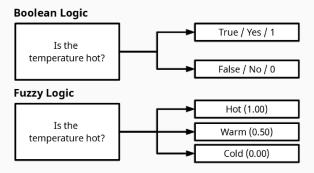
Fuzzy Logic, Microsoft® Encarta® Online Encyclopedia 2009 https://www.refseek.com/data/cache/en/1/Fuzzy_Logic.html

WHAT DOES FUZZY LOGIC HAVE TO OFFER?

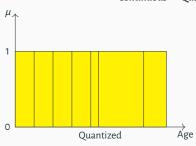
Fuzzy Logic aims at formalizing/mechanizing two noticeable human capabilities:

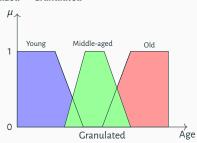
- 1. communicating, reasoning and rational decision making (in presence of imprecision, uncertainty & partiality of truth)
- 2. performing a wide variety of tasks (w/o measurements or computations)

FUZZY LOGIC AS AN EXTENSION OF THE BOOLEAN LOGIC

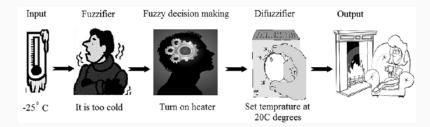


$\textbf{Continuous} \rightarrow \textbf{Quantized} \rightarrow \textbf{Granulated}$





EXAMPLE OF A FUZZY CONTROL SYSTEM



ARCHITECTURE

Rule Base is provided by experts. It contains the set of rules to govern the decision making.

Fuzzification converts crisp numbers to fuzzy sets.

Inference Engine decides which rules to be fired matching degree of the current fuzzy inputs.

Defuzzification converts the fuzzy sets delivered by the inference engine into some crisp value

28/79

DEFUZZIFICATION

A fuzzy value can be defuzzified through multiple ways.

- 1. Center of Sums
- 2. Centroid Method
- 3. Center of Area
- 4. Weighted Average Method
- 5. Max-Membership Principal

LOGICAL SYMBOLS

In formal logic \neg is <u>NOT</u>, \lor is **OR** and \land is **AND**

Consider the following propositions ${\cal E}$ and ${\cal S}$

 ${\mathcal E}$ "The earth is round"

 $\boldsymbol{\mathcal{S}}$ "The sun spins on its axis"

Thus

 $\neg \mathcal{E}$ "The earth is not round"

 $\neg S$ "The sun does not spin on its axis"

 $\mathcal{E} \vee \mathcal{S}$ "The earth is round **or** the sun spins on its axis"

 $\mathcal{E} \wedge \mathcal{S}$ "The earth is round **and** the sun spins on its axis"

 $\neg \mathcal{E} \lor \mathcal{S}$ "The earth is not round **or** the sun spins on its axis"

 $\mathcal{E} \wedge \neg \mathcal{S}$ "The earth is round **and** the sun does not spin on its axis"

CRISP VALUE IN FUZZY NOTATION (1/2)

Boolean Values as Fuzzy Sets

- ightharpoonup 0
 ightharpoonup 0 (completely non-membership)
- ► 1 → {1} (completely membership)
- ► In between 0 and 1, e.g., $0.5 \rightarrow \{0.5\}$ (partial membership)

Crisp Set Notation as Fuzzy Set Notation

A crisp set $\mathbf{A} = \{x \mid P(x)\}$ can be represented as a fuzzy set \mathcal{A} with membership function:

$$\mu_{\mathcal{R}}(x) = \begin{cases} 1 & \text{if } x \in \mathbf{A} \\ 0 & \text{if } x \notin \mathbf{A} \end{cases}$$

 \blacktriangleright Alternatively, a fuzzy set $\mathcal A$ can be defined with a membership function:

$$\mu_{\mathcal{A}}(x) = \begin{cases} 1 & \text{if } x \text{ is completely in } \mathcal{A} \\ 0.5 & \text{if } x \text{ is } \underline{\text{partially in }} \mathcal{A} \\ 0 & \text{if } x \text{ is completely out of } \mathcal{A} \end{cases}$$

31/79 A. MHAMDI Demystifying Al Sorcery

CRISP VALUE IN FUZZY NOTATION (2/2)

Fuzzy Notation for Crisp Set Elements

A crisp element $x \in \mathbf{U}$ can be represented as a fuzzy set $\{x\}$ with membership function:

$$\mu_{\{x\}}(x') = \delta(x, x')$$

where δ is the Dirac delta function.

Fuzzy Representation of Crisp Set Operations

For fuzzy sets \mathcal{A} and \mathcal{B} :

▶ Union $(\mathcal{A} \cup \mathcal{B})$:

$$\mu_{\mathcal{A} \cup \mathcal{B}}(x) = \max(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x))$$

▶ Intersection $(\mathcal{A} \cap \mathcal{B})$:

$$\mu_{\mathcal{A} \cap \mathcal{B}}(x) = \min(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x))$$

▶ Complement $(\neg \mathcal{A})$:

$$\mu_{\neg \mathcal{A}}(x) = 1 - \mu_{\mathcal{A}}(x)$$

Tipping Problem

What should be the TIP at a restaurant, given the quality of FOOD and of SERVICE. These latter are represented by some scores ranging from 0 (poor) to 10 (excellent).

Rules Base

- 1. FOOD is rancid ∨ SERVICE is poor ⇒ TIP is cheap;
- 2. SERVICE is good \Longrightarrow TIP is average;
- 3. FOOD is delicious ∨ SERVICE is excellent ⇒ TIP is generous

Tipping Problem

What should be the TIP at a restaurant, given the quality of FOOD and of SERVICE. These latter are represented by some scores ranging from 0 (poor) to 10 (excellent).

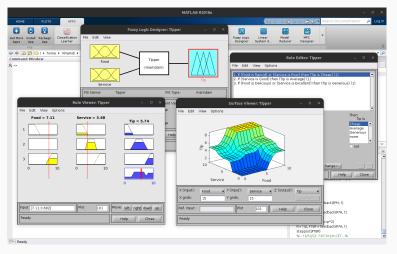
Rules Base

- 1. FOOD is rancid \lor SERVICE is poor \Longrightarrow TIP is cheap;
- 2. SERVICE is good \Longrightarrow TIP is average;
- 3. FOOD is delicious ∨ SERVICE is excellent ⇒ TIP is generous.

33/79

USING FUZZY LOGIC TOOLBOX



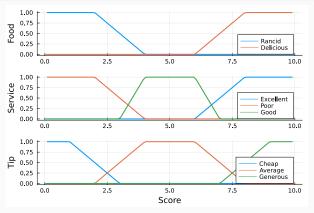




Code is available at https://github.com/a-mhamdi/cosnip/
→ Matlab → Fuzzy → Tipper fis

USING FUZZY.JL PACKAGE









FUZZY NUMBERS (1/6)

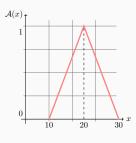
- ★ Represent imprecise numbers: number & linguistic modifier (e.g., nearly, around, etc.)
 - ► approximately five kilos
 - about 12 pm
- * Play an important role in decision making, approximate reasoning, statistics with imprecise probabilities and fuzzy control.

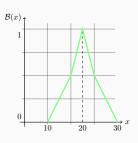
We need to perform arithmetic operations on fuzzy numbers (e.g., calculate a ratio of some fuzzy output over some fuzzy input)

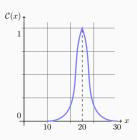
"around 20"

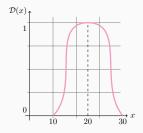
- ▶ includes some number values on either side of the central value of 20
- Central value is fully compatible with concept
- ▶ Number around central value are compatible with it to lesser degrees
- Degree of compatibility represented by fuzzy set; Membership value decreases from 1.0 to 0.0 on both sides of central value = fuzzy number.

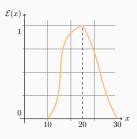
FUZZY NUMBERS (2/6)

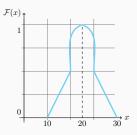












FUZZY NUMBERS (3/6)

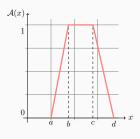
For a fuzzy membership function to qualify as a fuzzy number, it must capture our intuitive concept of a set of numbers around a given real number or interval of real numbers

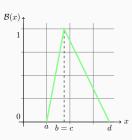
$$\mathcal{A}(x) = \begin{cases} f(x) & \text{for } x \in [a, b] \\ 1 & \text{for } x \in [b, c] \\ g(x) & \text{for } x \in [c, d] \\ 0 & \text{for } x < a \text{ or } x > d \end{cases}$$
 (1)

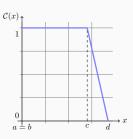
Common shapes of Fuzzy Numbers

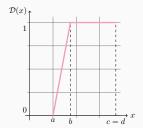
- Most common membership functions are trapezoidal and triangular (easy to construct and manipulate)
- ► Choice of a, b, c and d is important and is highly context-dependant
- ► Most applications not significantly affected by shapes of functions (i.e., use linear shapes)
- ▶ When some of real numbers (a, b, c, d) are equal, get degenerated forms of fuzzy numbers

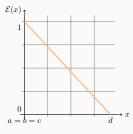
FUZZY NUMBERS (4/6)

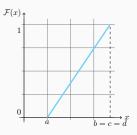






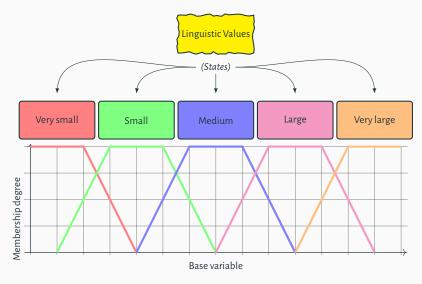






FUZZY NUMBERS (5/6)

States are fuzzy numbers which represent linguistic concepts



40/79 A. MH

FUZZY NUMBERS (6/6)

- 1. Fuzzy numbers are normal fuzzy sets (height=1)
- 2. Fuzzy numbers are convex fuzzy sets
- Support of every fuzzy number is open interval (a, d) of real numbers (support must be bounded)
- 4. Interval analysis can be used to define arithmetic operations on fuzzy numbers

Basic arithmetic operations:

- ightharpoonup addition [a, b] + [c, d] = [a + c, b + d]
- ► Subtraction [a, b] [c, d] = [a d, b c]
- ► Multiplication $[a, b] \times [c, d] = [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)]$
- ► Division² $[a, b] \div [c, d] = [a, b] \times [1/c, 1/d]$

Δ ΜΗΔΜΟΙ

 $^{^2}$ Interval division assumes that the number 0 is not one of the elements in the divisor interval [c, d].

CONSTRUCTING FUZZY SETS (1/5)



How would you assess today's temperature

We can describe a parameter describing a phenomena (e.g., Temperature for environment or Error for distance measurement) using a finite, small number of descriptors, referred to as linguistic variables of parameter.

Temperature (T) {Cold, Average, Warm}

Error (E) {Small, Medium, Large}



The number of linguistic variables should be kept small (7 \pm 2) due to our limited capacity to distinguish more. Commonly 3 to 5 linguistics variables are used in describing parameters.

CONSTRUCTING FUZZY SETS (2/5)



- ► Fuzzy sets offer an important and unique approach to describe linguistic variables
- ► Membership functions

$$\mathcal{A}(x) = X \rightarrow [0, 1]$$

are mathematical functions that are used to describe fuzzy sets

- ► Choosing membership functions require understanding of:
 - nature of the problem and parameter at hand
 - · Level of details to be captured
 - · Context of application

Prerequistes

- ► Concepts and linguistic values (e.g., cold temperature)
- ► Numerical measurements and/or linguistic assessments (e.g., degrees Celsius)
- Given context
- ▶ Data or Expert

CONSTRUCTING FUZZY SETS (3/5)

To construct fuzzy sets:

Expert-Driven Using developer, user, decision-maker, etc.

- 1. Direct methods
 - Answers to questions that explicitly pertain to the constructed membership function
 - · Single or multiple experts
- 2 indirect methods
 - Simpler questions, easier to answer, less sensitive to subjective biases, pertain to membership function only implicitly
 - Single or multiple experts

Data-Driven Form data to fuzzy sets

CONSTRUCTING FUZZY SETS (4/5)

Direct Methods with Multiple Experts

Example

n experts were asked to validate the proposition "x belongs to A" as either true or false

True
$$a_i(x) = 1$$

False
$$a_i(x) = 0$$

where $i \in \{1 \cdots n\}$ denotes the i^{th} expert.

$$A(x) = \frac{1}{n} \sum_{i=1}^{n} a_i(x)$$



Can also distinguish degrees of competence c_i of individual experts:

$$A(x) = \sum_{i=1}^{n} c_i a_i(x), \quad \text{where} \quad \sum_{i=1}^{n} c_i = 1$$

CONSTRUCTING FUZZY SETS (5/5)

| GIVEITS TABOUTETS (21, 22, 23, 24, 23) | · // · | |
|--|--------|---|
| Need to determine membership function "A" that | €#2 | 0 |
| captures linguistic term "Excellent Labourer" | €#3 | 0 |
| <u> </u> | Œ#4 | 1 |

- Ask 10 superintendents if particular person is excellent labourer (answer either yes (1) or no (0))
- For each labourer, calculate membership grade of belonging to fuzzy set "A" by taking ratio of total number of yes (1) to total number of responses.

. Civen 5 labourers (01 02 03 04 05)

| | £1 | \mathfrak{L} 2 | \mathfrak{L}_3 | $\mathfrak{L}4$ | \mathfrak{L}_{5} |
|-------------|----|------------------|------------------|-----------------|--------------------|
| E#1 | 1 | 1 | 1 | 1 | 1 |
| E#2 | 0 | 0 | 1 | 1 | 1 |
| €#3 | 0 | 1 | 0 | 1 | 0 |
| E#4 | 1 | 0 | 1 | 1 | 1 |
| E#5 | 0 | 0 | 1 | 1 | 1 |
| E# 6 | 0 | 1 | 1 | 1 | 1 |
| E#7 | 0 | 0 | 0 | 0 | 0 |
| E#8 | 1 | 1 | 1 | 1 | 1 |
| E#9 | 0 | 0 | 0 | 1 | 0 |
| €#10 | 0 | 0 | 0 | 1 | 0 |
| | | | | | |

⇒ Opinions of individual experts must be aggregated

The resulting set would be: $A = 0.3/\Omega + 0.4/\Omega + 0.6/\Omega + 0.6/\Omega + 0.6/\Omega$

FUZZY INFERENCE SYSTEMS (FIS)

MAMDANI ALGORITHM

A fuzzy logic controller that uses linguistic rules to map input values to output values.

- ▶ was introduced by Ebrahim (Abe) H. Mamdani in 1975
- works using rules of linguistics, style like human concepts (more intuitive and easier to understand)
- well suited to applications where rules are inspired from human expert knowledge.



We want to control the temperature of a living room using a thermostat.

- Inputs
- 1. ε difference between the desired and measured temperatures
- 2. $\delta \varepsilon$ rate of change of the temperature error

Output heating/cooling command(H/C) is the recommended action to take (e.g., heat, cool or no action)

FUZZY RULE BASE:

 \Re_1 : If ε is "cold" AND $\delta \varepsilon$ is "rapidly decreasing" THEN H/C is "heat"

 \Re_2 : If arepsilon is "hot" and \deltaarepsilon is "rapidly increasing" then H/C is "cool"

 \Re_3 : IF arepsilon is "near target" AND \deltaarepsilon is "slow" THEN H/C is "no action"

Task #1

Consider a fuzzy logic system with two inputs u, v and an output w. We suppose that each variable ranges from $0 \to 10$. w changes by a unit step. The membership functions of the fuzzy variables are described below.

u can be:

Negative (N) $\mathcal{L}(2, 4)$

Zero (Z) $\Delta(3, 6, 9)$

Positive (P) $\Gamma(6, 8)$.

▶ v can be-

Negative (N) $\mathcal{L}(2, 5)$

Zero (Z) $\Pi(2, 4, 6, 8)$

Positive (P) $\Gamma(6, 8)$.

w can be-

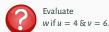
Small (S) $\mathcal{L}(2,4)$

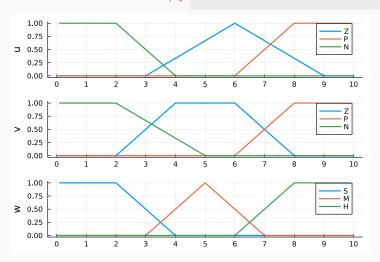
Medium (M) $\Delta(3, 5, 7)$

High (H) $\Gamma(6, 8)$.

Rule Base - case of ∧

| v | N | z | Р |
|---|---|---|---|
| N | S | S | М |
| Z | S | M | Н |
| Р | М | Н | Н |





$$\begin{cases}
4 = \{(\mathbf{N}, 0), (\mathbf{Z}, 1/3), (\mathbf{P}, 0)\} \\
& \Longrightarrow w^* = \frac{3 \times 0 + 4 \times 1/3 + 5 \times 1/3 + 6 \times 1/3 + 7 \times 0}{0 + 1/3 + 1/3 + 1/3 + 0} = 5
\end{cases}$$

Task #23

Design a fuzzy lighting controller system, in which the control system dims the bulb light automatically according to the environmental light. Assume that the inputs to the system are the environmental light x_1 and the changing rate of the environmental light x_2 . The output y represents the control value of the dimmer.

 $ightharpoonup x_1$ ranges between 120 and 220 lumens. x_1 can be:

Dark (D)
$$\mathcal{L}$$
 (130, 150)
Ambient (A) Π(130, 150, 190, 210)
Light (L) Γ(190, 210).

► x_2 ranges between -10 and +10. x_2 can be:

Negative-Small (NS)
$$\mathcal{L}(-10, 0)$$

Zero (Z) $\Delta(-10, 0, 10)$

Positive-Small (PS)
$$\Gamma(0, 10)$$
.

y ranges between 0 and +10. y can be:

Very-Small (VS)
$$\mathcal{L}(2, 4)$$

Small (S) $\Delta(2, 4, 6)$
Big (B) $\Delta(4, 6, 8)$

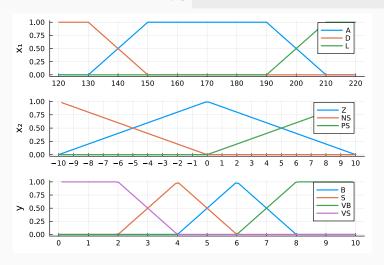
Very-Big (VB) $\Gamma(6, 8)$.

Rule Base - case of ∧

| x ₁ | D | Α | L |
|----------------|----|---|----|
| NS | VB | В | В |
| Z | В | В | S |
| PS | В | S | VS |



³Credit: Dr. Mohammed A. T.



125 = {(**D**, 1), (**A**, 0), (**L**, 0)} and -6 = {(**NS**, 3/5), (**Z**, 2/5), (**PS**, 0)}

$$\implies y^* = \frac{5 \times 2/5 + 6 \times 2/5 + 7 \times 1/2 + 8 \times 3/5 + 9 \times 3/5 + 10 \times 3/5}{2/5 + 2/5 + 3/5 + 1/2 + 3/5 + 3/5 + 3/5} \approx 7.775$$

TAKAGI-SUGENO-KANG (TSK)

Each fuzzy rule has the following structure: "IF (antecedent or premise) THEN (consequent)"

Antecedent composed by fuzzy sets defined by MF

Consequent represented by a polynomial function of the fuzzy inputs



We want to design a TSK FIS to control the temperature of a room. Our system accepts two inputs: room temperature (T) and outdoor temperature (T_0). The output is the heating or cooling action (H/C).

$$\Re_1$$
: If T is "cold" and T₀ is "cold" then H/C = $1/2$ T + $1/3$ T₀

$$\Re_2$$
: IF T is "cold" AND T₀ is "hot" THEN H/C = 1/5T + 3/5T₀

$$\Re_3$$
: IF T is "hot" AND T_o is "cold" THEN H/C = $-4/5T + 1/4T_0$

- ▲ Simple and efficient defuzzification
- ▲ Interpolation (incomplete or sparse rule bases)
- Smooth and continuous output
- ▼ Limited rule complexity
- ▼ Tuning (careful tuning of the fuzzy rule base)

Task#3

Suppose we have three fuzzy predicates: \mathcal{A} , \mathcal{B} and C described by these trapezoidal fuzzy sets:

x and y are fuzzy variables, each one ranges between 0 and 19. Given the following three rules:

$$\Re_1 (x \text{ is } \mathcal{A}) \land (y \text{ is } C) \rightarrow u = 10$$

 $\Re_2 \neg (x \text{ is } \mathcal{A}) \lor (y \text{ is } \mathcal{B}) \rightarrow u = 2$
 $\Re_3 (x \text{ is } \mathcal{B}) \land \neg (y \text{ is } C) \rightarrow u = 5$

Compute the degree of satisfaction for each case:

①
$$x_1 = 5 \& v_1 = 12$$
 ② $x_2 = 0 \& v_2 = 15$ ③ $x_3 = 7 \& v_3 = 13$

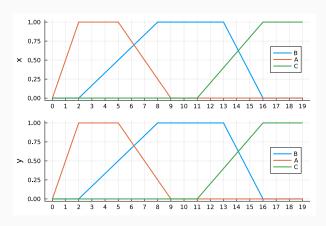
$$x_1 = 5 \& y_1 = 12$$

 \Re_1 $\mu_{\mathcal{A}}(x_1) \min \mu_{\mathcal{C}}(y_1)$ $1 \min \frac{1}{5} = \frac{1}{5}$

 \Re_2 $\mu_{\mathcal{A}}(x_1) \max \mu_{\mathcal{B}}(y_1)$ $0 \max 1 = 1$

 \Re_3 $\mu_{\mathcal{B}}(x_1) \min \mu_C(y_1)$ $1/2 \min 4/5 = 1/2$

 $u_1 = 3.82$



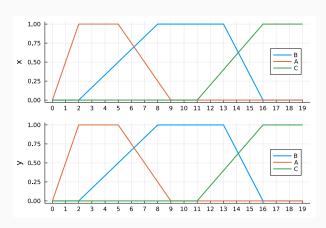
$$x_2 = 0 \& y_2 = 15$$

 \Re_1 $\mu_{\mathcal{A}}(x_2) \min \mu_{\mathcal{C}}(y_2)$ $0 \min _ = 0$

 \Re_2 $\mu_{\mathcal{A}}(x_2) \max \mu_{\mathcal{B}}(y_2)$ $1 \max 1/3 = 1$

 \Re_3 $\mu_{\mathcal{B}}(x_2) \min \mu_{\mathcal{C}}(y_2)$ $0 \min _ = 0$

 $u_2 = 2$



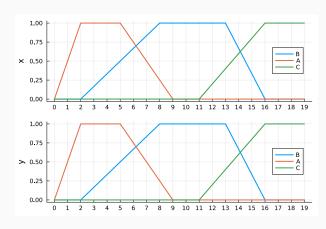
$$x_3 = 7 \& y_3 = 13$$

 \Re_1 $\mu_{\mathcal{A}}(x_3) \min \mu_C(y_3)$ $1/2 \min 2/5 = 2/5$

 \mathfrak{R}_2 $\mu_{\mathcal{A}}(x_3) \max \mu_{\mathcal{B}}(y_3)$ $1/2 \max 1 = 1$

 \Re_3 $\mu_{\mathcal{B}}(x_3) \min \mu_C(y_3)$ $\frac{5}{6} \min \frac{3}{5} = \frac{3}{5}$

 $u_3 = 4.5$



Task #44

Let us consider a TSK fuzzy model with the input (x_1, x_2) and a single output y. Evaluate y^* for $(x_1, x_2) = (1.5, 0.5)$. The rules are:

- 1. If x_1 is S AND x_2 is S THEN $y = -x_1 + x_2 + 1$
- 2. IF x_1 is S AND x_2 is \mathcal{L} THEN $y = -x_2 + 3$
- 3. IF x_1 is \mathcal{L} AND x_2 is \mathcal{S} THEN $y = -x_1 + 3$
- 4. IF x_1 is \mathcal{L} AND x_2 is \mathcal{L} THEN $y = -x_1 + x_2 + 2$

The fuzzy predicates ${\cal S}$ and ${\cal L}$ are described by these membership functions:

$$S \ \mathcal{L}(-2, 2)$$

⁴Credit: Dr. Hashim A. H.

$$x_1 = 1.5 \& x_2 = 0.5$$

 \Re_1

$$\mu_{S}(x_{1}) \min \mu_{S}(x_{2})$$
0.125 min 0.375 = 0.125
 $(y = 0)$

 \Re_2

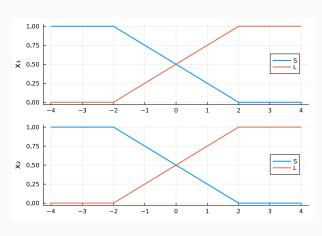
$$\mu_{\mathcal{S}}(x_1) \min \mu_{\mathcal{L}}(x_2)$$
0.125 min 0.625 = 0.125
 $(y = 2.5)$

Ú

$$\Re_3$$
 $\mu_{\mathcal{L}}(x_1) \min \mu_{\mathcal{S}}(x_2)$
0.875 min 0.375 = 0.375
 $(y = 1.5)$

 \Re_4

$$\mu_{\mathcal{L}}(x_1) \min \mu_{\mathcal{L}}(x_2)$$
0.875 min 0.625 = 0.625
 $(y = 1)$



$$y^* = \frac{0 \times 0.125 + 2.5 \times 0.125 + 1.5 \times 0.375 + 1 \times 0.625}{0.125 + 0.125 + 0.375 + 0.625} = 1.2$$

TSUKAMOTO

A type of fuzzy logic-based system developed by Tsukamoto in the 1970s. The system would compute the crisp output for each rule based on the firing strength and then take the weighted average to produce the final performance evaluation.

EXAMPLE ____

We want to evaluate employee performance based on two criteria:

1 job satisfaction and 2 productivity

We can define three fuzzy sets for each criterion: "low" (L), "medium" (M), and "high" (H)

FUZZY RULE BASE:

 $\mathfrak{R}_1\colon$ IF job satisfaction is L AND productivity is L, THEN performance is L.

 \Re_2 : IF job satisfaction is **M** AND productivity is **M**, THEN performance is **M**.

 $\mathfrak{R}_3\colon$ If job satisfaction is H and productivity is H, then performance is H.

▲ Simplified Rule Base

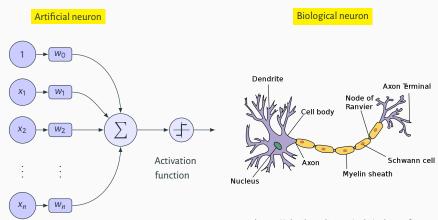
Faster Computation

▼ Limited Transparency

Limited Flexibility

Neural Networks

FUNDAMENTAL UNIT OF A NEURAL NETWORK (1/3)



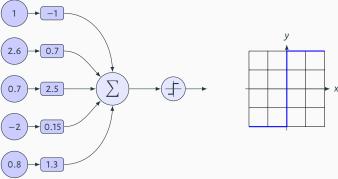
Inputs Weights

https://id.wikipedia.org/wiki/Sel_saraf

FUNDAMENTAL UNIT OF A NEURAL NETWORK (2/3)

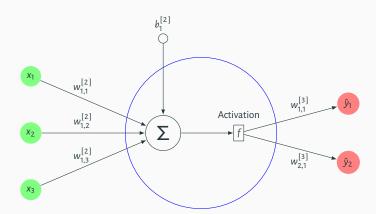
Task #5

Compute the output of the following neuron.

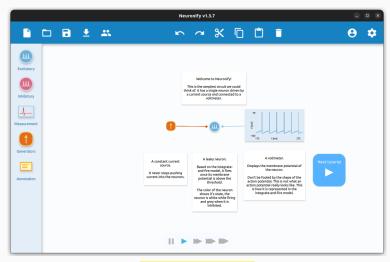


$$y = sign(1 \times -1 + 2.6 \times 0.7 + 0.7 \times 2.5 - 2 \times 0.15 + 0.8 \times 1.3) = 1$$

FUNDAMENTAL UNIT OF A NEURAL NETWORK (3/3)

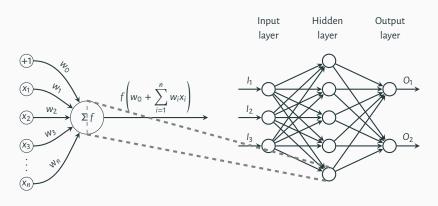


NEURAL SIMULATION



http://ovilab.net/neuronify/

MULTILAYER PERCEPTRON (MLP)



Task #6
For the above structure, determine how many parameters are to be adjusted.

$$\# params = 5 \times 3 + 5 + 2 \times 5 + 2 = 32$$

- ✓ Design a structure
- ✓ Specify a loss function to minimize
- ✓ Optimize using gradient descent
 - ① Feedforward propagation (matrix multiplication and point-wise activation)
 - ② Back propagation (multivariate chain rule)
 - ③ Update the weights accordingly

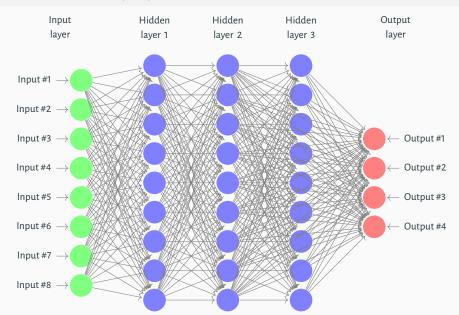
 Input

 Synaptic Weights

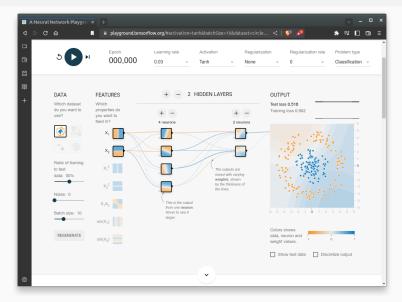
 Compare

 UPDATE WEIGHTS

MULTILAYER PERCEPTRON (MLP)

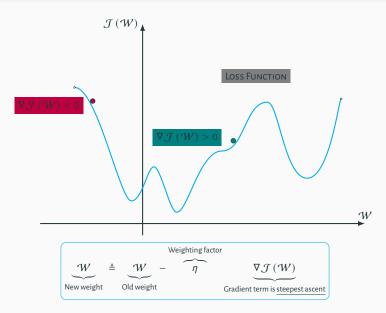


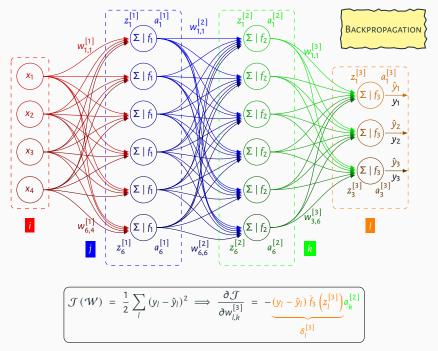
TINKER WITH A NEURAL NETWORK



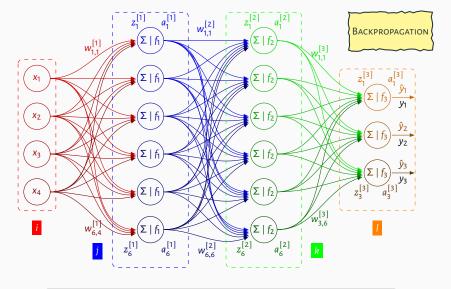
https://playground.tensorflow.org/

GRADIENT DESCENT



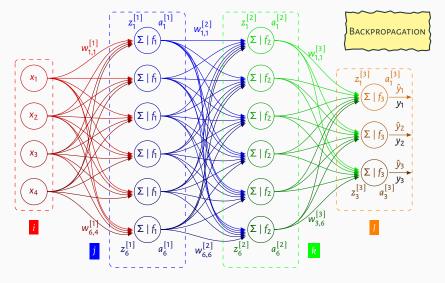


67/79 A. MHAMDI Demystifying Al Sorcery

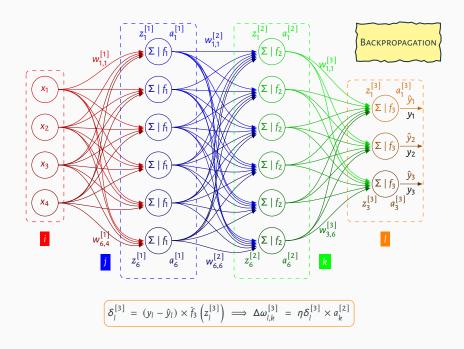


$$\frac{\partial \mathcal{J}}{\partial w_{k,j}^{[2]}} = -\sum_{l} (y_{l} - \hat{y}_{l}) \dot{f}_{3} \left(z_{l}^{[3]} \right) w_{l,k}^{[3]} \dot{f}_{2} \left(z_{k}^{[2]} \right) a_{j}^{[1]} = -\underbrace{\sum_{l} \delta_{l}^{[3]} w_{l,k}^{[3]} \dot{f}_{2} \left(z_{k}^{[2]} \right) a_{j}^{[1]}}_{\delta_{k}^{[2]}}$$

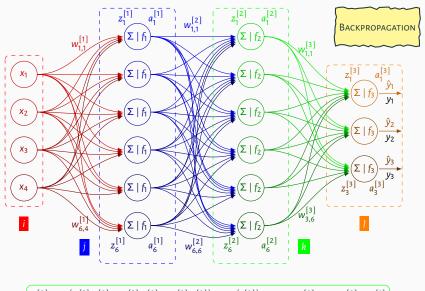
67/79



$$\left(\frac{\partial \mathcal{J}}{\partial w_{j,i}^{[1]}} = -\sum_{l} (y_{l} - \hat{y}_{l}) \dot{f}_{3} \left(z_{l}^{[3]}\right) \sum_{k} w_{l,k}^{[3]} \dot{f}_{2} \left(z_{k}^{[2]}\right) w_{k,j}^{[2]} \dot{f}_{1} \left(z_{j}^{[1]}\right) x_{i} = -\underbrace{\sum_{k} \delta_{k}^{[2]} w_{k,j}^{[2]} \dot{f}_{1} \left(z_{j}^{[1]}\right) x_{i}}_{\delta_{j}^{[1]}}\right) \left(\frac{\partial \mathcal{J}}{\partial w_{j,i}^{[1]}}\right) \left(\frac{\partial \mathcal{J}}{\partial w_{j,$$

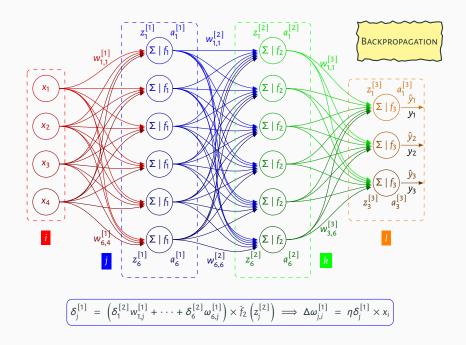


67/79 A. Mhamdi Demystifying Al Sorcery



$$\label{eq:delta_k} \boxed{ \delta_k^{[2]} \ = \ \left(\delta_1^{[3]} w_{1,k}^{[2]} + \delta_2^{[3]} \omega_{2,k}^{[2]} + \delta_3^{[3]} \omega_{3,k}^{[2]} \right) \\ \times \dot{f}_2 \left(z_k^{[2]} \right) \ \Longrightarrow \ \Delta \omega_{k,j}^{[2]} \ = \ \eta \delta_k^{[2]} \times a_j^{[1]} } }$$

67/79 A. MHAMDI Demystifying Al Sorcery



MULTIVARIATE CHAIN RULE

Output layer → hidden layer #2

$$\frac{\partial \hat{y}_l}{\partial w_{l,k}^{[3]}} = \underbrace{\frac{\partial \hat{y}_l}{\partial z_l^{[3]}}}_{j_3\left(z_l^{[3]}\right)} \underbrace{\frac{\partial z_l^{[3]}}{\partial w_{l,k}^{[3]}}}_{a_k^{[2]}}$$

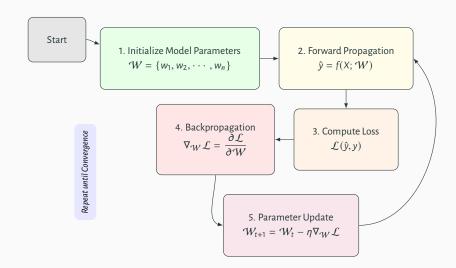
Output layer → hidden layer #1

$$\frac{\partial \hat{y}_{l}}{\partial w_{k,j}^{[2]}} = \underbrace{\frac{\partial \hat{y}_{l}}{\partial z_{l}^{[3]}}}_{\hat{f}_{3}(z_{l}^{[3]})} \underbrace{\frac{\partial z_{l}^{[3]}}{\partial a_{k}^{[2]}}}_{w_{l,k}^{[3]}} \underbrace{\frac{\partial z_{l}^{[2]}}{\partial z_{k}^{[2]}}}_{\hat{f}_{2}(z_{k}^{[2]})} \underbrace{\frac{\partial z_{l}^{[2]}}{\partial w_{k,j}^{[1]}}}_{a_{l}^{[1]}}$$

Output layer → input layer

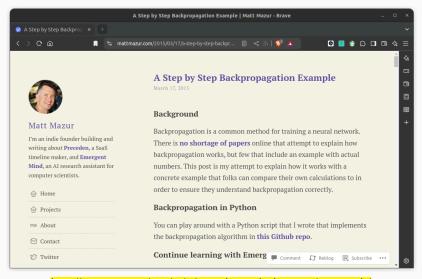
$$\frac{\partial \hat{y}_{l}}{\partial w_{j,i}^{[1]}} = \underbrace{\frac{\partial \hat{y}_{l}}{\partial z_{l}^{[3]}}}_{f_{3}(z_{l}^{[3]})} \underbrace{\frac{\partial z_{l}^{[3]}}{\partial a_{k}^{[2]}}}_{g_{k}^{[2]}} \underbrace{\frac{\partial z_{k}^{[2]}}{\partial z_{k}^{[2]}}}_{g_{k}^{[1]}} \underbrace{\frac{\partial z_{l}^{[1]}}{\partial z_{l}^{[1]}}}_{g_{k}^{[1]}} \underbrace{\frac{\partial z_{l}^{[1]}}{\partial z_{j}^{[1]}}}_{g_{k}^{[1]}} \underbrace{\frac{\partial z_{l}^{[1]}}{\partial w_{j,i}^{[1]}}}_{x_{l}}$$

NEURAL NETWORK TRAINING



69/79

BACKPROPAGATION BY HAND



https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/

70/79 A. MHAMDI Demystifying Al Sorcery

LIST OF AVAILABLE OPTIMIZERS (1/2)

These are a few typical optimizers for artificial neural networks:

$$\boxed{ \Delta \hat{\mathcal{W}} \triangleq \mathcal{F} \left(\nabla \underbrace{\mathcal{J} \left(\hat{\mathcal{W}} \right)}_{\text{Loss Function}} \right) \equiv \underbrace{ \hat{\mathcal{W}} \triangleq \hat{\mathcal{W}} + \mathcal{F} \left(\nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \right) }_{\text{Loss Function}} \nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) = \begin{bmatrix} \frac{\partial \mathcal{J}}{\partial \hat{w}_0} \\ \vdots \\ \frac{\partial \mathcal{J}}{\partial \hat{w}_n} \end{bmatrix}$$

Stochastic Gradient Descent (SGD)

$$\hat{W} \triangleq \hat{W} - \eta \nabla \mathcal{J}(\hat{W})$$

Mini-batch Gradient Descent

$$\hat{W} \triangleq \hat{W} - \frac{\eta}{m} \nabla \sum_{i=1}^{m} \mathcal{J}(\hat{W}) \qquad \longleftarrow m \text{ denotes the size of the mini-batch}$$

Momentum

$$\hat{W} \triangleq \hat{W} - \mathcal{V}$$
, where $\mathcal{V} \triangleq \alpha \mathcal{V} + \eta \nabla \mathcal{J} (\hat{W})$

Nesterov Accelerated Gradient (NAG)

$$\hat{W} \triangleq \hat{W} - \mathcal{V}$$
 where $\mathcal{V} \triangleq \alpha \mathcal{V} + \eta \nabla \mathcal{J} (\hat{W} - \alpha \mathcal{V})$

LIST OF AVAILABLE OPTIMIZERS (2/2)

AdaGrad

$$\hat{\mathcal{W}} \triangleq \hat{\mathcal{W}} - \frac{\eta}{\sqrt{\mathcal{G} + \epsilon}} \nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \quad \text{where} \quad \mathcal{G} \triangleq \mathcal{G} + \left(\nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \right)^2$$

RMSProp

$$\hat{\mathcal{W}} \triangleq \hat{\mathcal{W}} - \frac{\eta}{\sqrt{\mathcal{G} + \epsilon}} \nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \quad \text{where} \quad \mathcal{G} \triangleq \mathcal{G} + (1 - \beta) \left(\nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \right)^2$$

Adam

$$\mathcal{M} \triangleq \beta_1 \mathcal{M} + (1 - \beta_1) \nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \qquad \longleftarrow \text{Estimate of first moment}$$

$$\mathcal{V} \triangleq \beta_2 \mathcal{V} + (1 - \beta_2) \left(\nabla \mathcal{J} \left(\hat{\mathcal{W}} \right) \right)^2 \qquad \longleftarrow \text{Estimate of second moment}$$

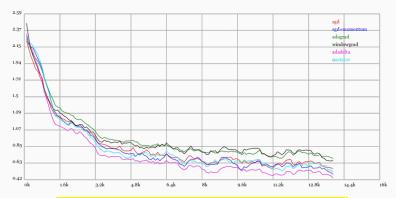
$$\hat{\mathcal{M}} = \frac{\mathcal{M}}{1 - \beta_1^k} \qquad \longleftarrow \text{@ every } k^{\text{th}} \text{ iteration}$$

$$\hat{\mathcal{V}} = \frac{\mathcal{V}}{1 - \beta_2^k} \qquad \longleftarrow \text{@ every } k^{\text{th}} \text{ iteration}$$

$$\hat{\mathcal{W}} \triangleq \hat{\mathcal{W}} - \frac{\eta}{\sqrt{\hat{\mathcal{V}} + \delta}} \hat{\mathcal{M}}$$

72/79

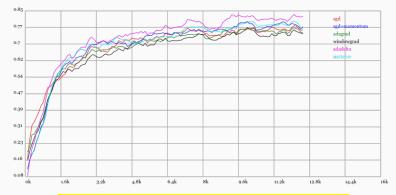
EFFECT OF OPTIMIZER ON LOSS VALUES



https://cs.stanford.edu/people/karpathy/convnetjs/demo/trainers.html

73/79 A. MHAMDI Demystifying Al Sorcery

EFFECT OF OPTIMIZER ON TESTING ACCURACY VALUES



https://cs.stanford.edu/people/karpathy/convnetjs/demo/trainers.html

74/79 A. MHAMDI Demystifying Al Sorcery

FRAMEWORKS TO BE USED





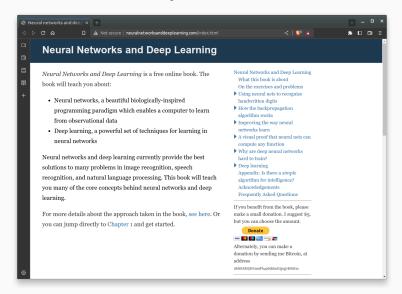


CODE SNIPPET





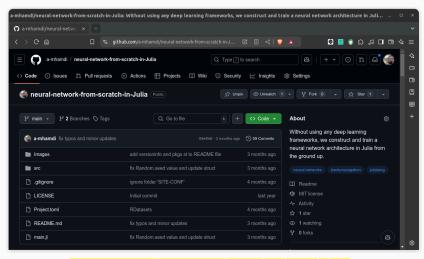
► This will be continued in the following e-book.



http://neuralnetworksanddeeplearning.com/

NEURAL NETWORK FROM SCRATCH

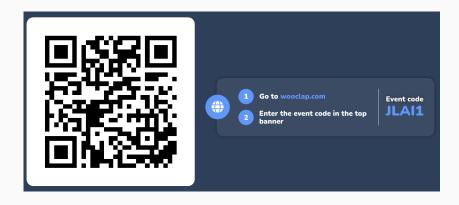




https://github.com/a-mhamdi/neural-network-from-scratch-in-Julia

Quizzes

KNOWLEDGE CHECK



https://app.wooclap.com/JLAI1

FURTHER READING (1/3)

References

[Bel78] R. E. Bellman, An Introduction to Artificial Intelligence: Can Computers Think? Boyd & Fraser Publishing Company, Jan. 1, 1978 (cit. on p. 11). [CMM85] E. Charniak, D. McDermott, and D. V. McDermott. Introduction to Artificial Intelligence. Addison-Wesley series in computer science and information processing. Addison-Wesley, 1985 (cit. on p. 12). [Dad12] E. Dadios, ed. Fuzzy Logic - Controls, Concepts, Theories and Applications. IntechOpen, Mar. 28, 2012. 430 pp. [ENM15] I. El Naga and M. J. Murphy. "What Is Machine Learning?" In: Machine Learning in Radiation Oncology: Theory and Applications. Ed. by I. El Naga, R. Li, and M. J. Murphy. Cham: Springer International Publishing, 2015, pp. 3–11. DOI: 10.1007/978-3-319-18305-3_1. L. Gacôgne. Intelligence artificielle. Cours. exercices corrigés et projets. Ed. by Ellipses. Nov. [Gac15] 2015. 240 pp. [GBC16] I. Goodfellow, J. Bengio, and A. Courville. Deep Learning. MIT Press Ltd., Nov. 18, 2016. 800 pp. [Hau89]]. Haugeland. Artificial Intelligence: The Very Idea. A Bradford book, MIT Press, 1989 (cit. on p. 11).

FURTHER READING (2/3)

- [JPM21] L. M. John Paul Mueller. Machine Learning For Dummies. Wiley John + Sons, Apr. 8, 2021.
 464 pp.
- [Kur92] R. Kurzweil. *The Age of Intelligent Machines*. Viking, 1992 (cit. on p. 13).
- [LS93] G. F. Luger and W. A. Stubblefield. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Artificial intelligence. Benjamin/Cummings Publishing Company, 1993 (cit. on p. 14).
- [Mit97] T. M. Mitchell. *Machine Learning*. McGraw-Hill International Editions. McGraw-Hill, 1997.
- [RK91] E. Rich and K. Knight. Artificial Intelligence. Artificial Intelligence Series. McGraw-Hill, 1991 (cit. on p. 13).
- [Rob14] A. Robinson. *Construction Informatics*. Ed. by H. S. of Construction Engineering. 2014.
- [Sch90] R. J. Schalkoff. Artificial Intelligence: An Engineering Approach. McGraw-Hill Computer science series. McGraw-Hill, 1990 (cit. on p. 14).
- [SNK12] T. Sai, D. Nakhaeinia, and B. Karasfi. "Application of Fuzzy Logic in Mobile Robot Navigation". In: Fuzzy Logic - Controls, Concepts, Theories and Applications. InTech, Mar. 2012. DOI: 10.5772/36358.
- [Win92] P. H. Winston. Artificial Intelligence. A-W Series in Computer Science. Addison-Wesley Publishing Company, 1992 (cit. on p. 12).

FURTHER READING (3/3)

[Woj12] J. Wojtusiak. "Machine Learning". In: Encyclopedia of the Sciences of Learning. Springer US, 2012, pp. 2082–2083. DOI: 10.1007/978-1-4419-1428-6_1927.