

Lesson 1: Agent

by Hand 

Prof. Tom Yeh



University of Colorado
Boulder

Hosted by



Roadmap

Q/A

LLM

RAG

Agent

Assignment

Patterns



Money

AutoGen



Equation

See

Think

Remember

Can



Manus

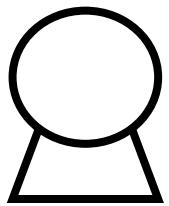
LLM -> RAG -> Agent

Agent - AI by Hand 🖊

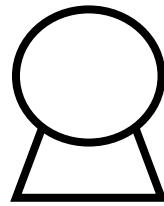


University of Colorado
Boulder

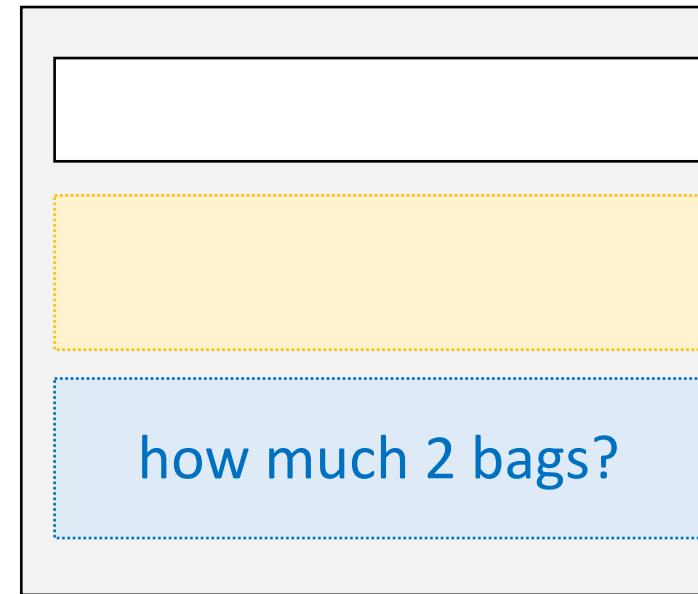
LLM Only



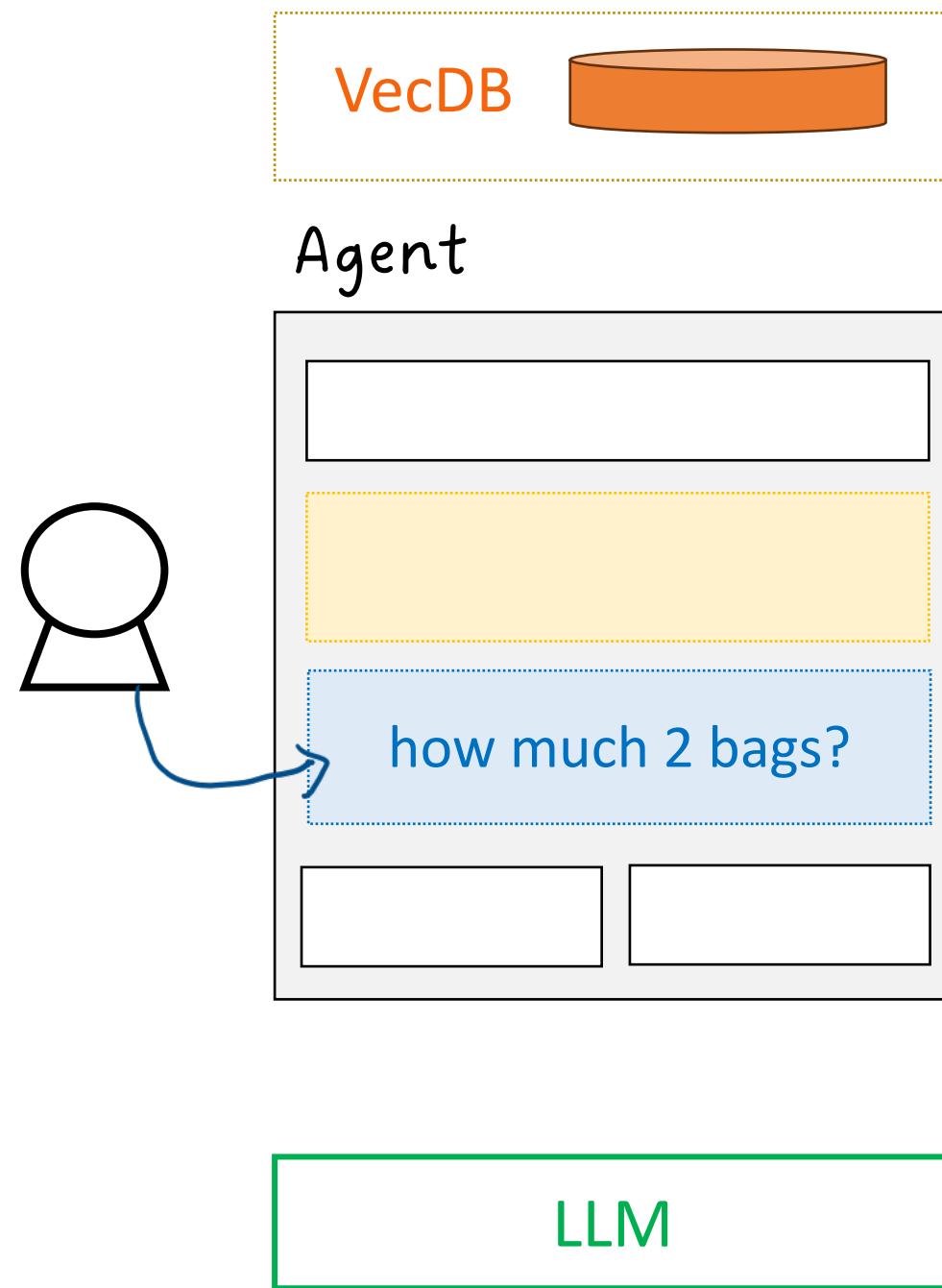
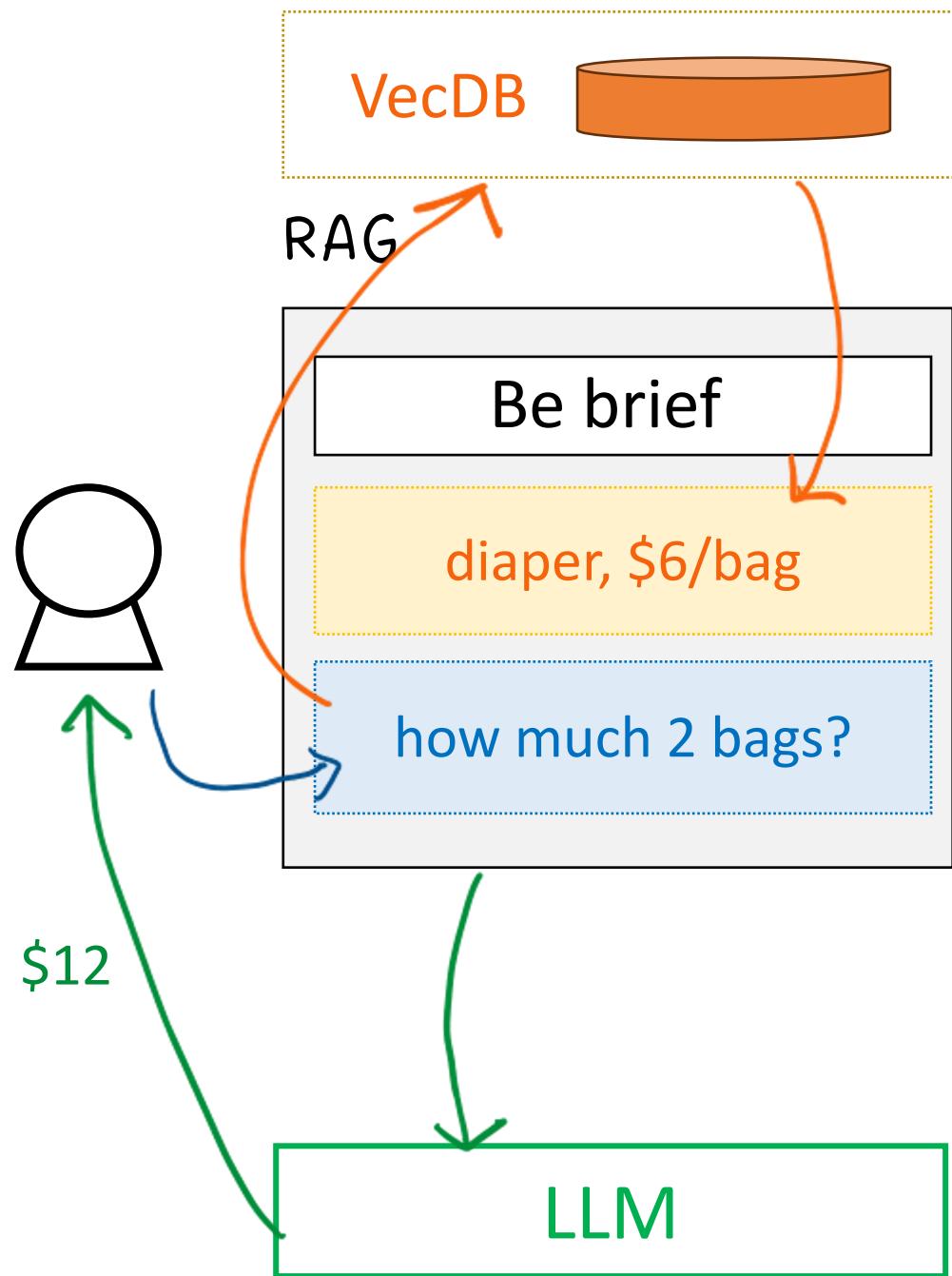
RAG



RAG



LLM



AI by Hand around the World

Agent - AI by Hand 🖍



University of Colorado
Boulder

Name: Pustam

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

Country: Egypt

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

Rim Missadui

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

Country: Viet Nam

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

India

2	1	3
-1	3	5
1	5	3

$\phi: \text{ReLU}$

1	-1	1	-5
1	1	0	0
0	1	1	1
1	0	1	-2

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Problem:

1	-1	1	5
1	1	0	0
0	1	1	1
1	0	1	-2

$\phi: \text{ReLU}$

2	1	3
-1	3	5
1	5	3

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Output

L_1

L_2

$\phi: \text{ReLU}$

INDIA

2	1	3
-1	3	5
1	5	3

$\phi: \text{ReLU}$

1	-1	1	-5
1	1	0	0
0	1	1	1
1	0	1	-2

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Honor Roll

AI by Hand

27 people from 19 countries

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

Name: Abhinav Shukla

Country: India

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

AI By Hand © 2024 Tom Yeh

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

1	0	1	0
-1	0	1	0
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

3	1	1	1
1	1	0	1
-1	0	1	0
0	1	1	-3

\approx

3	1	1	1
1	1	0	1
0	1	1	-3
1	1	0	0

$\phi: \text{ReLU}$

CHALLENGE:

2	1	3
-1	3	5
1	5	3

$\phi: \text{ReLU}$

1	-1	1	-5
1	1	0	0
0	1	1	1
1	0	1	-2

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Input

2	1	3
-1	3	5
1	5	3

$\phi: \text{ReLU}$

1	-1	1	-5
1	1	0	0
0	1	1	1
1	0	1	-2

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Input →

2	1	3
-1	3	5
1	5	3

$\phi: \text{ReLU}$

1	-1	1	-5
1	1	0	0
0	1	1	1
1	0	1	-2

\approx

0	3	5
3	5	3
3	3	3

$\phi: \text{ReLU}$

Neural Network

AI by Hand 🖐

Tom

Sudhin

Bhishan



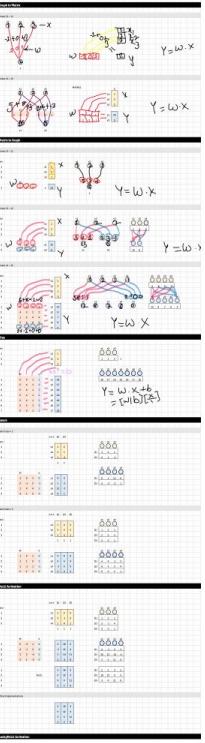
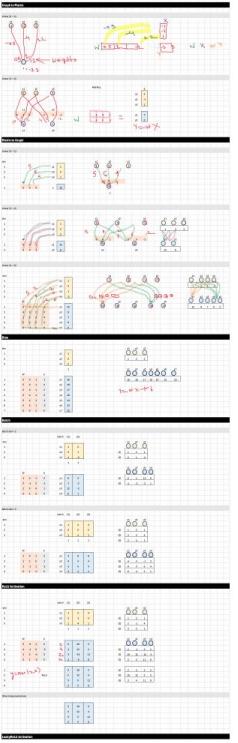
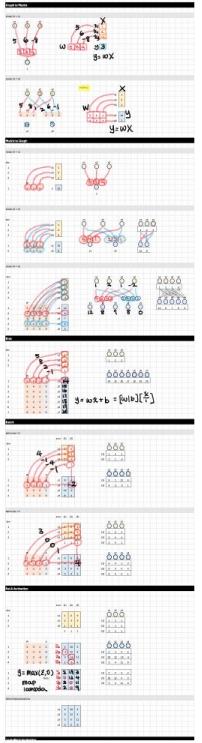
Boulder
USA



Pokhar
Nepal

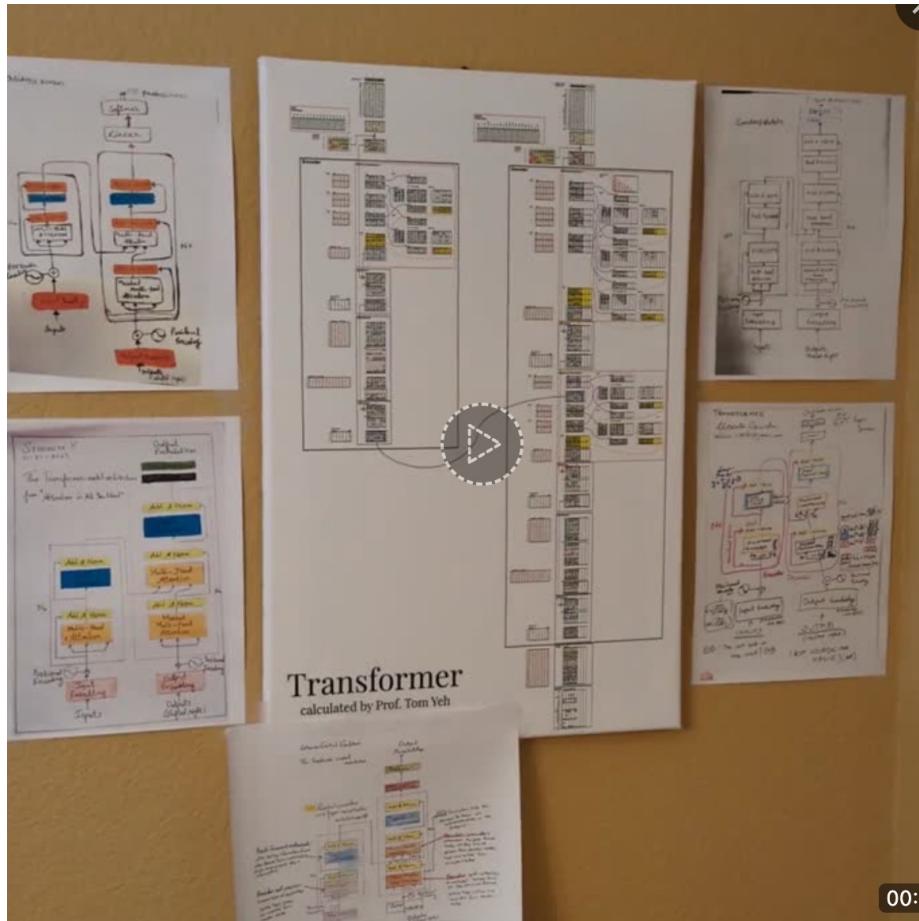


Kathmandu
Nepal



Watch my lecture videos at

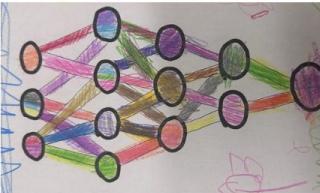
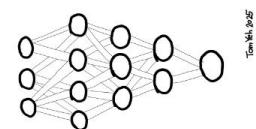
byhand.ai/youtube



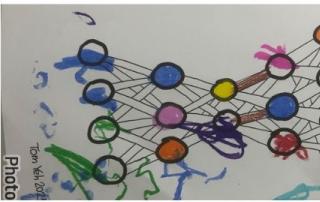
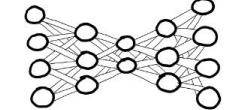
Color AI by Hand 🖐 Togeth

see my post to participate in the group art project

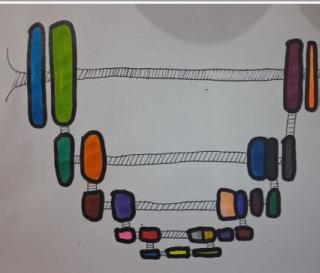
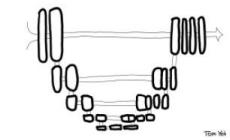
Neural Net



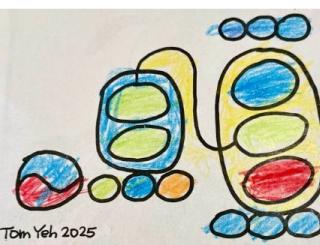
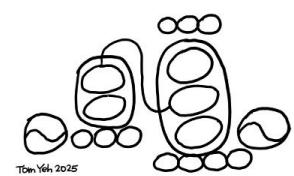
Auto-Encoder



U-Net



Transformer



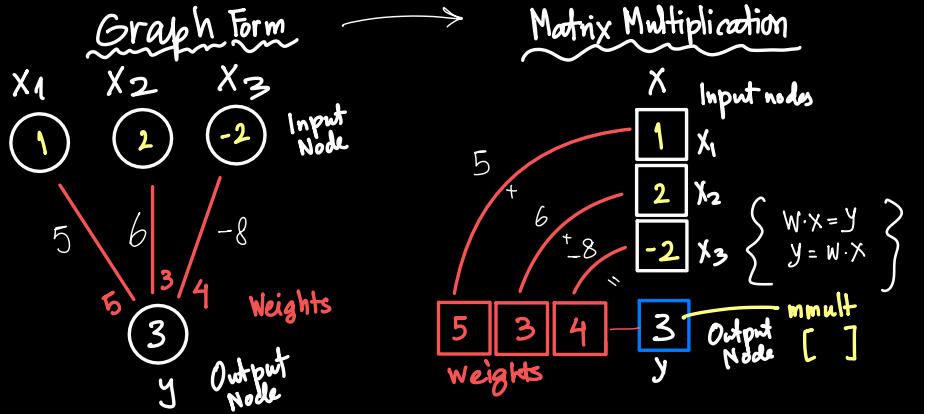
Graphical representation of Neural Network

mathematical underpinning of the BlackBox

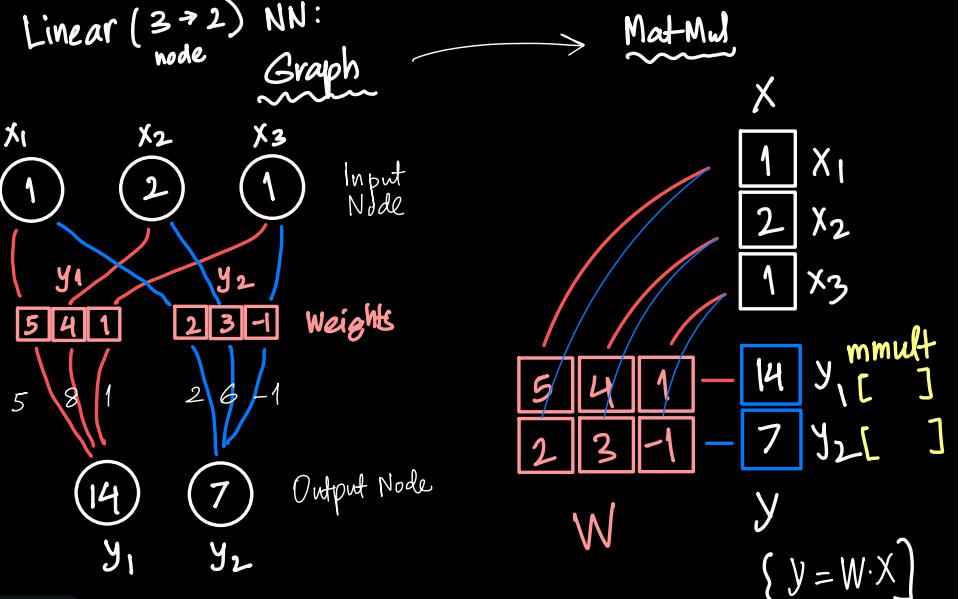
/@ajunkodher
1

10

* Linear Layer : goes from $(3 \rightarrow 1)$ Nodes



* Linear $(3 \rightarrow 2)$ node NN:



Made with Goodnotes

* ReLU

h_1	0	0	0	1	3
h_2	9	10	8	6	1
1	1	1	1	1	1

$$\max(\boxed{\quad}, 0) \quad \text{ReLU}(w_1 x_1 + w_2 x_2 + b)$$

0	0	3	6	9	12	15	18	21	24	27
0	0	2	5	8	11	14	17	20	23	26
0	0	1	4	7	10	13	16	19	22	25
0	0	0	3	6	9	12	15	18	21	24
1	0	0	0	2	5	8	11	14	17	20
0	0	0	1	4	7	10	13	16	19	22
-1	0	0	0	0	3	6	9	12	15	18
-2	0	0	0	0	2	5	8	11	14	17
-3	0	0	0	0	1	4	7	10	13	16
-4	0	0	0	0	0	3	6	9	12	15
-5	0	0	0	0	0	2	5	8	11	14

$$\max(\boxed{\quad}, 0) \quad \text{ReLU}(w_1 x_1 + w_2 x_2 + b)$$

27	25	23	21	19	17	15	13	11	9	7
24	22	20	18	16	14	12	10	8	6	4
21	19	17	15	13	11	9	7	5	3	1
18	16	14	12	10	8	6	4	2	0	0
15	13	11	9	7	5	3	1	0	0	0
12	10	8	6	4	2	0	0	0	0	0
-1	9	7	5	3	0	0	0	0	0	0
-2	6	4	2	0	0	0	0	0	0	0
-3	3	1	0	0	0	0	0	0	0	0
-4	0	0	0	0	0	0	0	0	0	0
-5	0	0	0	0	0	0	0	0	0	0

Output Layer

$$\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$$

$w_1 \quad w_2 \quad b$

$$w_1^{\text{(1st hidden node)}} + w_2^{\text{(2nd hidden node)}} + b \rightarrow$$

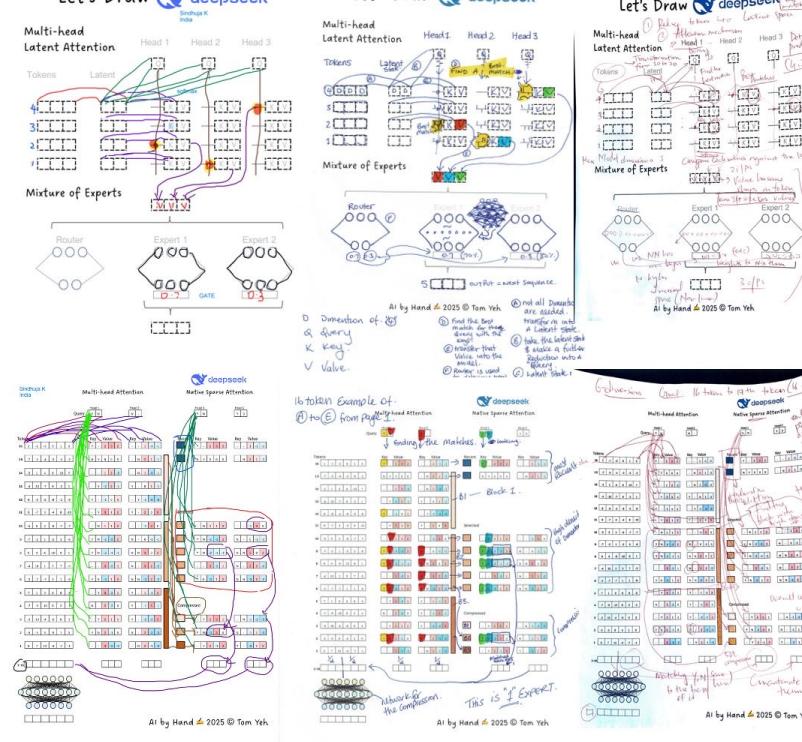
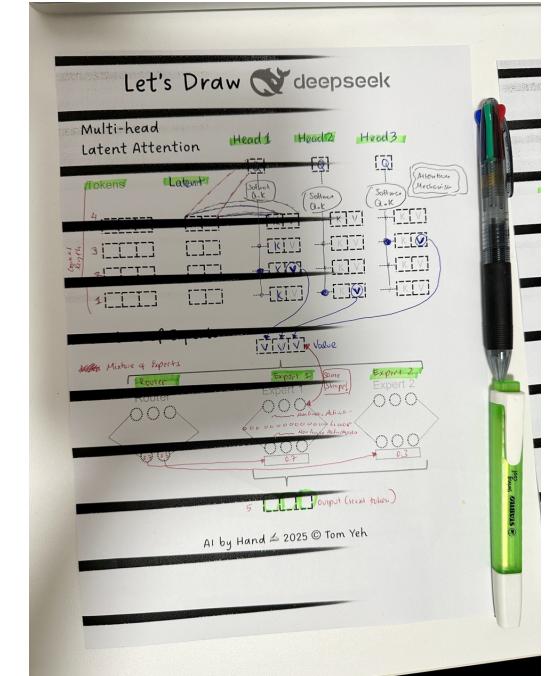
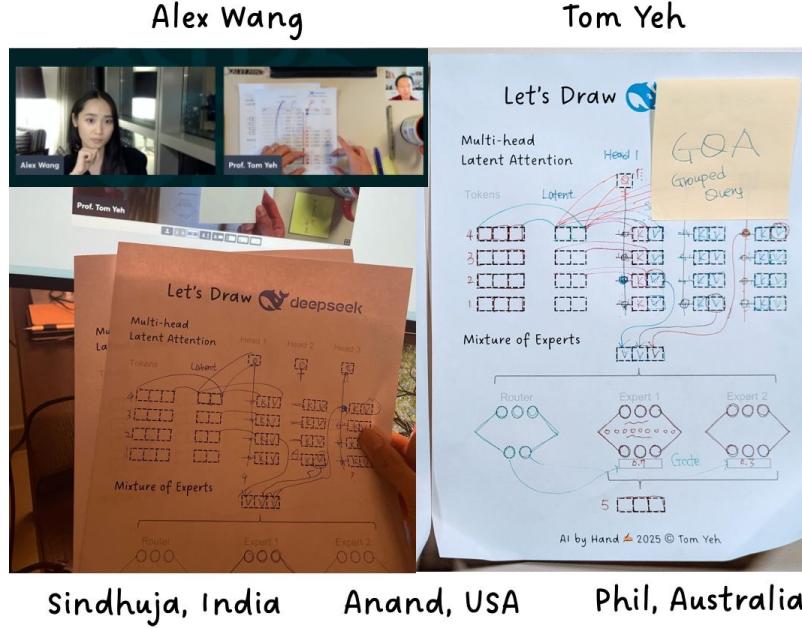
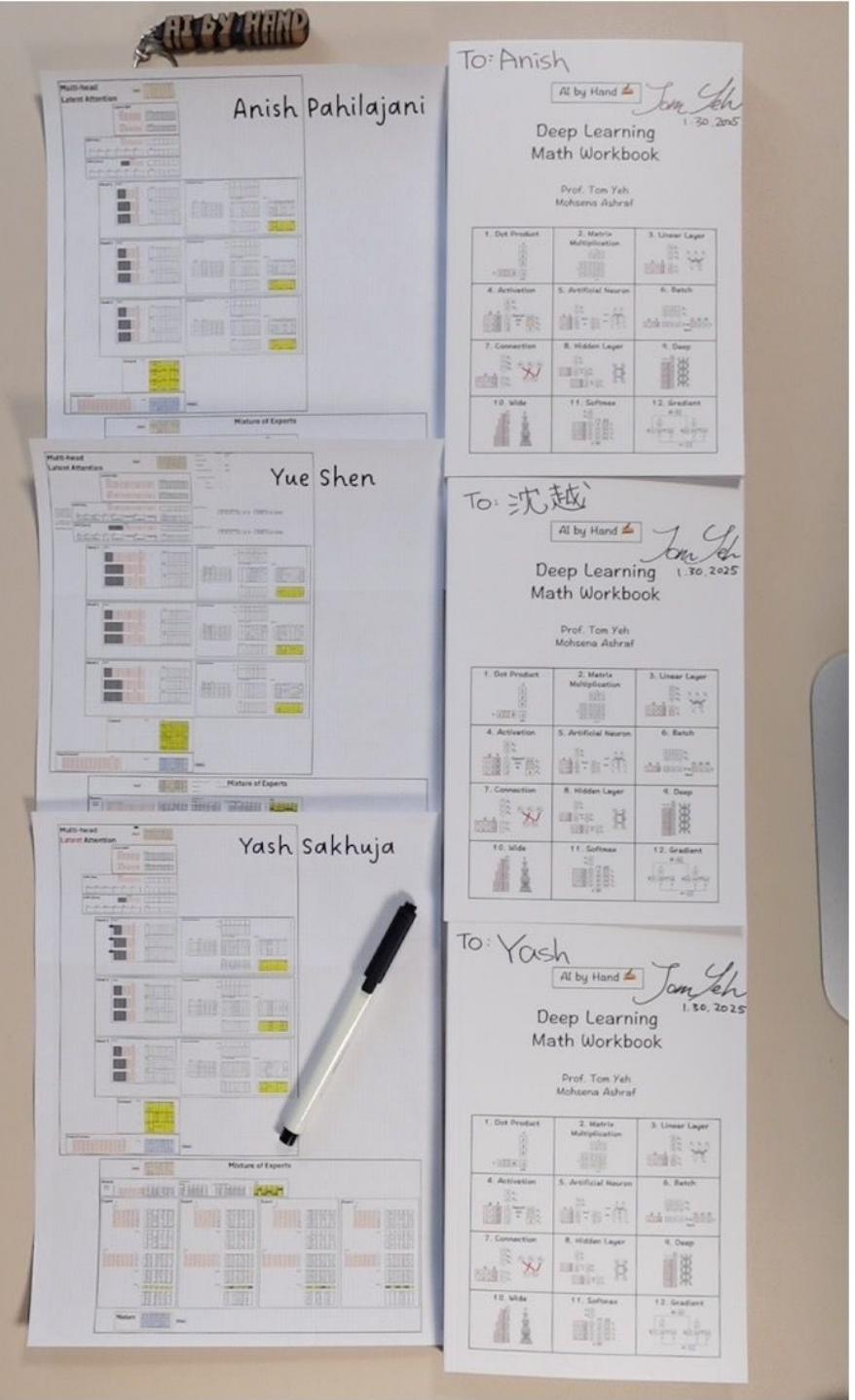
Sigmoid:

$$P \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} = 1 / (1 + \text{EXP}(-\boxed{\quad}))$$

decision boundary

1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	1
1	0.9	1	1	0.9	1	1	1	1	1	1
2	0.9	1	0.7	0.3	0.7	1	1	1	1	1
3	0.8	0.5	0.3	0.3	0.5	1	1	1	1	1
4	0.26	0.3	0.3	0.3	0.3	0.9	1	1	1	1
5	0.26	0.3	0.3	0.3	0.3	0.7	1	1	1	1

10



Certificate

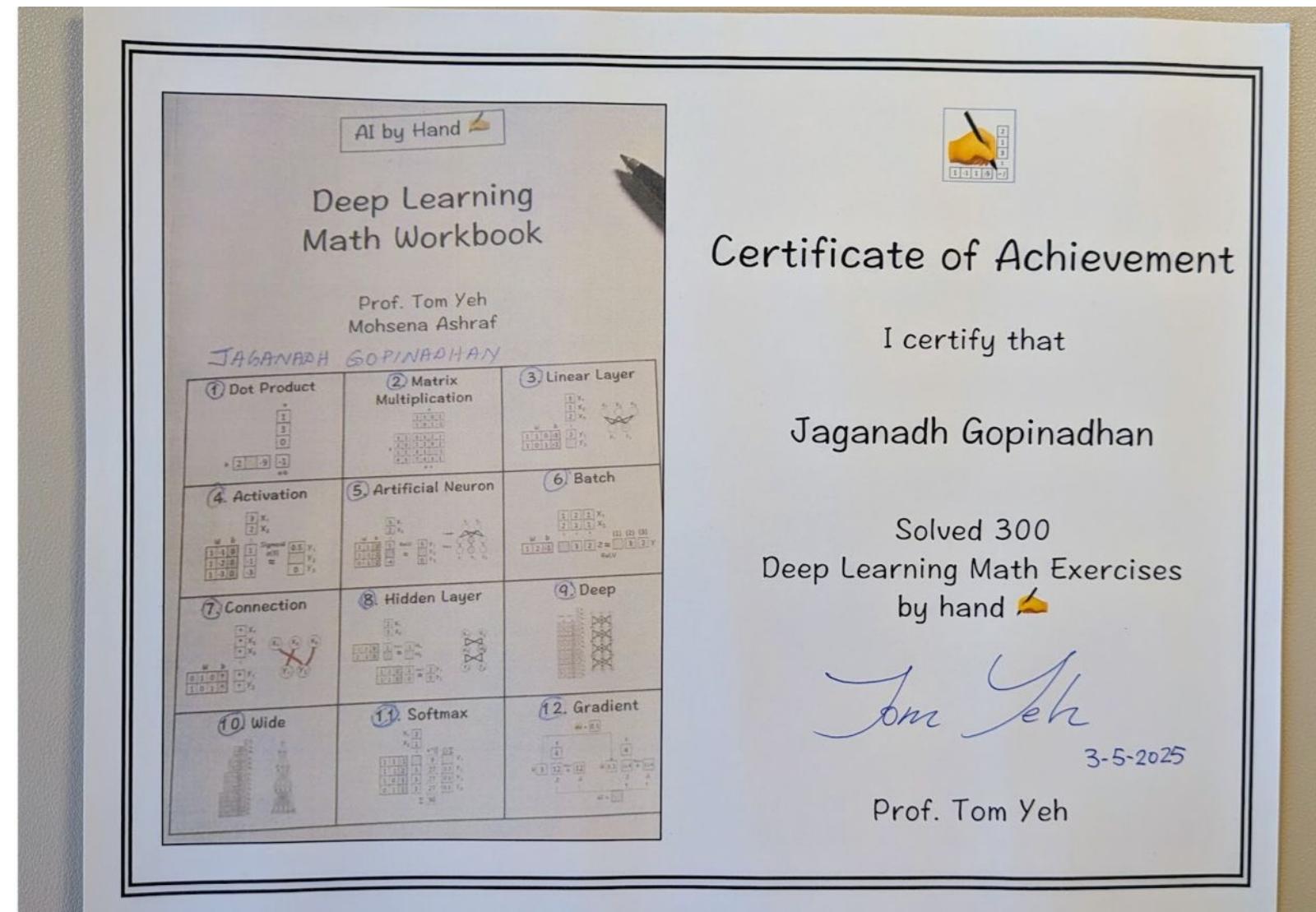
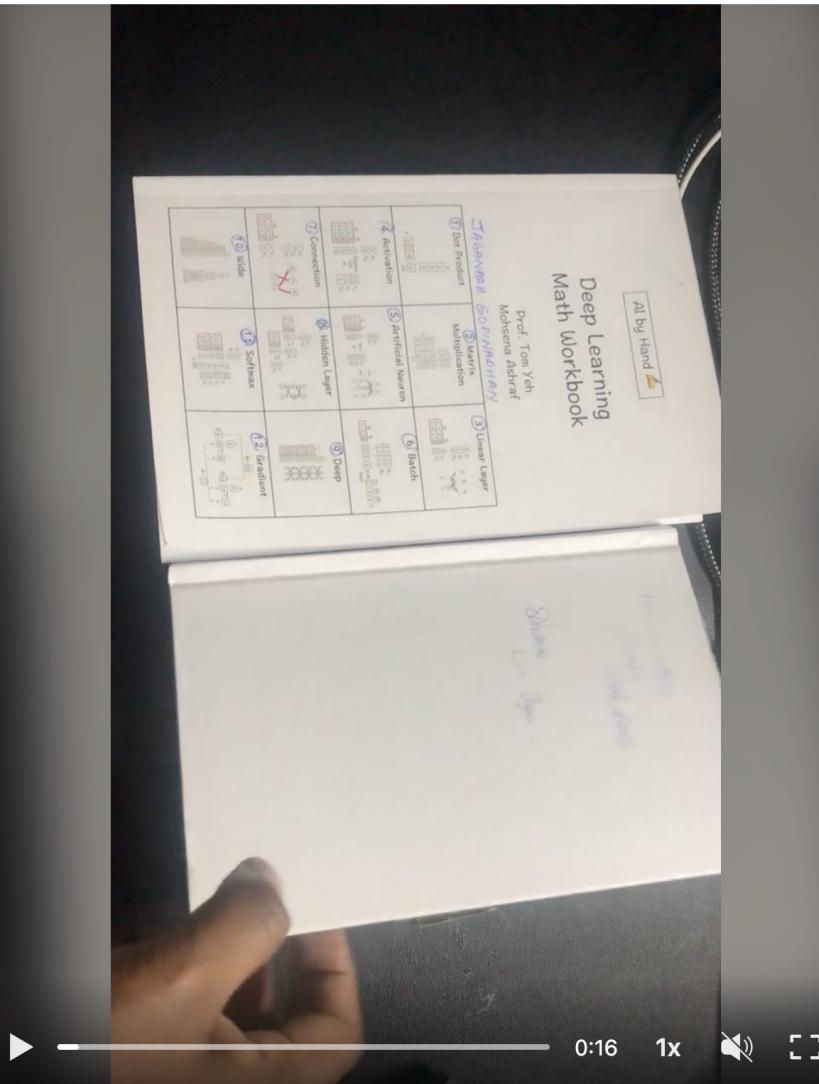
Agent - AI by Hand 🖊



University of Colorado
Boulder

Finally, I completed all 12 chapters from the AI by Hand Deep Learning Math Workbook. Solving each of the problems was a great experience. Thanks to **Tom Yeh** and **Mohsena Ashraf** for creating such excellent material.
I'm looking forward to other workbooks, too.

#aibyhand



Patterns

Agent - AI by Hand 🖍



University of Colorado
Boulder

LLM Only

Prompt

Summarize this article:
Quantum physics, also known as quantum mechanics, is a fundamental theory in physics that describes nature at the smallest scales of energy levels of atoms and subatomic particles. Central to the theory is the concept that energy, typically thought of as continuously...

LLM

Certainly! Here's the summary:

Quantum physics explores subatomic particles' wave-particle duality and uncertainty.

Agent Pattern 1: Reflection

Prompt

Summarize this article:
Quantum physics, also known as quantum mechanics, is a fundamental theory in physics that describes nature at the smallest scales of energy levels of atoms and subatomic particles. Central to the theory is the concept that energy, typically thought of as continuously...

LLM

Certainly! Here's the summary:

Quantum physics explores subatomic particles' wave-particle duality and uncertainty.

Agent

Do I need to know more about this request?

Agent Pattern 1: Reflection

Prompt

Summarize this article:
Quantum physics, also known as quantum mechanics, is a fundamental theory in physics that describes nature at the smallest scales of energy levels of atoms and subatomic particles. Central to the theory is the concept that energy, typically thought of as continuously...

LLM

Certainly! Here's the summary:

Quantum physics explores subatomic particles' wave-particle duality and uncertainty.

Agent

Should I really do this?

Agent Pattern 2: Tool Use

Prompt

What is the airfare from LA
to NYC?

LLM

\$500

Agent

Should I check an
external source?

Agent Pattern 2: Tool Use

Prompt

What is the airfare from LA
to NYC for my family of
five?

LLM

\$2500

Agent

Should I use an
external app?

Agent Pattern 3: Planning

Prompt

Summarize this article:
Quantum physics, also known as quantum mechanics, is a fundamental theory in physics that describes nature at the smallest scales of energy levels of atoms and subatomic particles. Central to the theory is the concept that energy, typically thought of as continuously...

LLM

Certainly! Here's the summary:

Quantum physics explores subatomic particles' wave-particle duality and uncertainty.

Agent

What is the first step?

Agent Pattern 4: Multi-Agent Planning

Prompt

Can you give me a discount?
NYC?

LLM

No.

Agent

Who else can do
this better?

Recognize LLM Agent Patterns

User: What is the meaning of life?

It's 42.	<input type="radio"/>
I will use a search engine.	<input type="radio"/>
First, understand the motivation.	<input type="radio"/>
I will ask a Librarian.	<input type="radio"/>
Please clarify. Whose life?	<input type="radio"/>

<input type="radio"/>	LLM only
<input type="radio"/>	Reflection
<input type="radio"/>	Tool Use
<input type="radio"/>	Planning
<input type="radio"/>	Multi Agent

Money

Agent - AI by Hand 🖊



University of Colorado
Boulder

How much does an LLM cost?

User Query

What is the airfare
from LA to NYC?

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

LLM

Five hundred
dollars

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

Total: \$_____

How much does an Agent cost?

Use Query

What is the airfare
from LA to NYC?

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

Agent Prompt

You are a helpful
assistant
You can use Google

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

Agent Response

Five hundred
dollars

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

Total: \$_____

Agent Equation

Agent - AI by Hand 🖊



University of Colorado
Boulder

Equation

Agent = L 

+ S 

+ T 

+ R 

+ C 

Think, See, Remember, Can

Agent - AI by Hand 🖍



University of Colorado
Boulder

Example: Buy a house

User: Please help me buy a house.

See

You saw

Generate a response.

Think

You saw

<user> help me buy a house

You think

Generate a response.

Remember

You saw

<user> help me buy a house

You think

<role> a helpful realtor

<plan> step 1 - know your client

You remembered

Generate a response.

Can

You saw

<user> help me buy a house

You think

<role> a helpful realtor

<plan> step 1 - know your client

You remembered

<history> I helped rent a place last year.

<data> Three new houses are just listed.

You can

Generate a response.

Complete Prompt

You saw

<user> help me buy a house

You think

<role> a helpful realtor

<plan> step 1 - know your client

You remembered

<history> I helped rent a place last year.

<data> Three new houses are just listed.

You can

1. Talk to the seller agent.
2. Book an appointment.

Generate a response.

{ prompt | completion}
tokens x _____

\$_____ / token

Cost: \$_____

Example: Test a program

User: Please test this program l1m.cpp

You saw

Prompt tokens x _____

You think

You remembered

You can

Generate a response.



The general AI agent

Case Study

Agent - AI by Hand 🖌



University of Colorado
Boulder



F Forbes

China's Autonomous Agent, Manus, Changes Everything

1 day ago

ET The Economic Times

Another Chinese company rattles Silicon Valley, Manus AI stuns tech world with...



2 days ago

bt Business Today

'Outperforming DeepResearch': New Chinese AI agent Manus rewrites...



10 hours ago

Ecns.cn

China's new AI contender Manus stirs another buzz



South China Morning Post

Chinese AI agent Manus faces doubts after rapid rise to fame

2 days ago

TC TechCrunch

Manus probably isn't China's second 'DeepSeek moment'



2 hours ago

in India Today

After DeepSeek, Manus AI comes out of China: What is it and why is it making a...



12 hours ago

IE Interesting Engineering

'World's first' fully autonomous AI agent unveiled in China, handles...



8 hours ago

Nikkei Asia

The next DeepSeek? Startup Manus becomes China's latest AI darling



2 days ago

Use case gallery

Learn how Manus handles real-world tasks through step-by-step replays.

Featured

Research

Life

Data Analysis

Education

Productivity

WTF



Interactive Course on the Momentum Theorem

Manus develops engaging video presentations for middle school educators, clearly explaining the momentum theorem through accessible



Campaign Explanation Maps

Custom-designed visualization maps that bring historical events like the Battle of Lexington to life, enhancing student understanding



Interactive Transformer Course

Elegantly designed interactive webpages that demystify Transformer architecture through clear explanations and engaging visual elements.



Exploring the Scale of the Universe

Interactive interactive course



Converting Lecture Recordings into Notes

Transform your classroom



Learning Resource Collection and Organization

You think, saw, remembered, can

manus Interactive Webpage for Learning Trans... Log in

Design an interactive webpage that explains the Transformer architecture with clear visuals and engaging elements. Structure it as a step-by-step learning journey from fundamentals to advanced concepts, incorporating interactive demonstrations that showcase the Transformer's capabilities and elegance.

Manus

I'll help you design an interactive webpage that explains the Transformer architecture with clear visuals and engaging elements. I'll create a step-by-step learning journey that covers everything from fundamentals to advanced concepts, with interactive demonstrations to showcase the Transformer's capabilities. Let me get started on this right away!

Suggested new knowledge

- ✓ Research Transformer architecture comprehensively ^
Creating project structure and preparing to research Transformer architecture for the interactive webpage.
Creating file transformer_webpage/todo.md
- ✓ Create todo file and project structure ^
- ✓ Research Transformer architecture comprehensively ^
Researching Transformer architecture fundamentals and attention mechanisms from multiple sources.
Browsing https://www.machinelearningmastery.com/the-transformer-atten...

Scrolling down

Manus's Computer

Manus is using Browser
Scrolling down

https://machinelearningmastery.co...

Let's get started.

Buiding Transformer Models with Attention Mechanisms: Build a Neural Machine Translator in 12 Days
Adding a Custom Attention Layer to a Recurrent Neural Network in Keras
A Gentle Introduction to Positional Encoding in Transformer Models, Part 1
Training the Transformer Model
Joining the Transformer Encoder and Decoder Parts
Meeting

The Transformer attention mechanism
Photo by Arunesh Choudhury, some rights reserved.

Tutorial Overview
This tutorial is divided into two parts; they are:

- Introduction to the Transformer Attention
- The Transformer Architecture
 - Scaled-Dot Product Attention
 - Multi-Head Attention

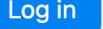
Prerequisites
For this tutorial, we assume that you are already familiar with:

- The concept of attention

Manus is working: Implement detailed explanations

You think, saw, remembered, can

 manus Interactive Webpage for Learning Transf...



 Manus

I'll help you design an interactive webpage that explains the Transformer architecture with clear visuals and engaging elements. I'll create a step-by-step learning journey that covers everything from fundamentals to advanced concepts, with interactive demonstrations to showcase the Transformer's capabilities. Let me get started on this right away!

 Suggested new knowledge

✓ Research Transformer architecture comprehensively ^
Creating project structure and preparing to research Transformer architecture for the interactive webpage.
 Creating file transformer_webpage/todo.md

✓ Create todo file and project structure ^
Project structure created with detailed todo list. Moving to research phase on Transformer architecture.
 Searching transformer architecture deep learning attention mechanism

✓ Research Transformer architecture comprehensively ^
Researching Transformer architecture fundamentals and attention mechanisms from multiple sources.
 Browsing <https://www.machinelearningmastery.com/the-transformer-attent...>
Researching Transformer architecture fundamentals by exploring information about attention mechanisms from multiple sources.
 Scrolling down

Manus's Computer

Manus is using Editor
Editing file `transformer_webpage/todo.md`

`todo.md`

Interactive Transformer Diff Original Modified

Research Phase

- [x] Research Transformer architecture fundamentals
- [x] Gather information on attention mechanisms
- [x] Research multi-head attention
- [x] Study positional encoding
- [x] Research feed-forward networks in Transformers
- [x] Gather information on encoder-decoder structure
- [x] Research applications of Transformers
- [] Collect visualization ideas for Transformer components

Design Phase

- [x] Design webpage layout
- [x] Plan step-by-step learning journey
- [x] Design visual elements for each component
- [x] Plan interactive demonstrations
- [x] Create wireframes for the webpage

Development Phase

- [] Set up project structure
- [] Choose appropriate web framework (Next.js)
- [] Implement base layout and navigation
- [] Develop visualizations for attention mechanism
- [] Create visualizations for multi-head attention
- [] Implement positional encoding visualization
- [] Develop encoder-decoder structure



AutoGen

Agent - AI by Hand 🖌



University of Colorado
Boulder

AutoGen

```
agent = ConversableAgent(  
    _____,  
    system_message=_____,  
    llm_config={"config_list": [  
        {"model": _____,  
         "temperature": _____,  
         "api_key": _____} ]}  
)
```

AutoGen

```
agent = ConversableAgent(  
    "cathy",  
    system_message="Your are Cathy, a part of a duo of comedians.",  
    llm_config={"config_list": [  
        {"model": "gpt-4",  
         "temperature": 0.9,  
         "api_key": OPENAI_API_KEY} ]}  
)
```

Default System Message

DEFAULT_SYSTEM_MESSAGE = """You are a helpful AI assistant.

Solve tasks using your coding and language skills.

In the following cases, suggest python code (in a python coding block) or shell script (in a sh coding block) for the user to execute.

1. When you need to collect info, use the code to output the info you need, for example, browse or search the web, download/read a file, print the content of a webpage or a file, get the current date/time, check the operating system. After sufficient info is printed and the task is ready to be solved based on your language skill, you can solve the task by yourself.

2. When you need to perform some task with code, use the code to perform the task and output the result. Finish the task smartly.

Solve the task step by step if you need to. If a plan is not provided, explain your plan first. Be clear which step uses code, and which step uses your language skill.

When using code, you must indicate the script type in the code block. The user cannot provide any other feedback or perform any other action beyond executing the code you suggest. The user can't modify your code. So do not suggest incomplete code which requires users to modify. Don't use a code block if it's not intended to be executed by the user. If you want the user to save the code in a file before executing it, put # filename: <filename> inside the code block as the first line. Don't include multiple code blocks in one response. Do not ask users to copy and paste the result. Instead, use 'print' function for the output when relevant. Check the execution result returned by the user.

If the result indicates there is an error, fix the error and output the code again. Suggest the full code instead of partial code or code changes. If the error can't be fixed or if the task is not solved even after the code is executed successfully, analyze the problem, revisit your assumption, collect additional info you need, and think of a different approach to try.

When you find an answer, verify the answer carefully. Include verifiable evidence in your response if possible.

Reply "TERMINATE" in the end when everything is done.

"""

You think, saw, remembered, can

DEFAULT_SYSTEM_MESSAGE = """You are a helpful AI assistant.

Solve tasks using your coding and language skills.

In the following cases, suggest python code (in a python coding block) or shell script (in a sh coding block) for the user to execute.

1. When you need to collect info, use the code to output the info you need, for example, browse or search the web, download/read a file, print the content of a webpage or a file, get the current date/time, check the operating system. After sufficient info is printed and the task is ready to be solved based on your language skill, you can solve the task by yourself.

2. When you need to perform some task with code, use the code to perform the task and output the result. Finish the task smartly.

Solve the task step by step if you need to. If a plan is not provided, explain your plan first. Be clear which step uses code, and which step uses your language skill.

...Do not ask users to copy and paste the result. Instead, use 'print' function for the output when relevant. Check the execution result returned by the user.

If the result indicates there is an error, fix the error and output the code again. Suggest the full code instead of partial code or code changes. If the error can't be fixed or if the task is not solved even after the code is executed successfully, analyze the problem, revisit your assumption, collect additional info you need, and think of a different approach to try.

When you find an answer, verify the answer carefully. Include verifiable evidence in your response if possible.

Reply "TERMINATE" in the end when everything is done."""



crewAI

Agent - AI by Hand



University of Colorado
Boulder

crewAI

```
writer = Agent(  
    role='Writer',  
    goal='Narrate compelling tech stories about {topic}',  
    memory=True,  
    backstory=  
        "You craft engaging narratives to educate"  
    ),  
    tools=[search_tool],  
    allow_delegation=True  
)
```

You think, saw, remembered, can

```
writer = Agent(  
    role='Writer',  
    goal='Narrate compelling tech stories about {topic}',  
    memory=True,  
    backstory=  
        "You craft engaging narratives to educate"  
    ),  
    tools=[search_tool],  
    allow_delegation=True  
)
```

Assignment

Agent - AI by Hand 🖊



University of Colorado
Boulder

Prizes

Agent - AI by Hand 🖍



University of Colorado
Boulder

Q/A

Agent - AI by Hand 🖊



University of Colorado
Boulder