(Gentle) Introduction to Modern AI, Transformers and Agentic solutions

יוסי זגורי
Yossi@integrative.net
052-4668866

Topics:
AI
DL
TRANSFORMERS
LLM
OLLAMA

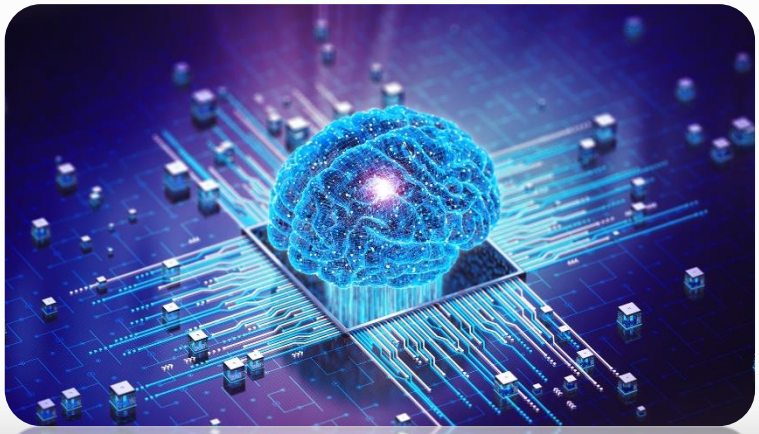
AI Background

Cognition, Intelligence & AI

Cognition refers to the human mental processes involved in acquiring knowledge and understanding through thought, experience, and senses, using perception, memory, attention, reasoning, problem-solving, and decision-making

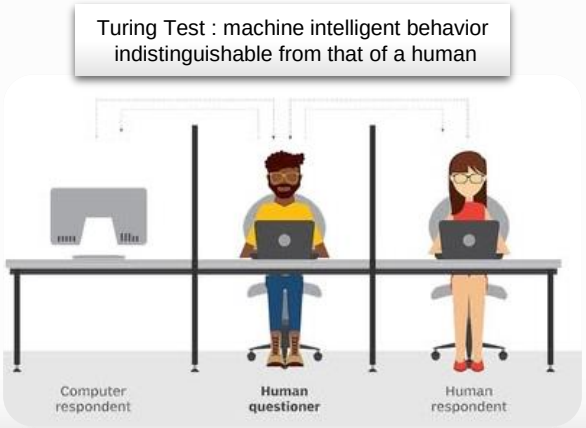
Intelligence is often considered the ability to use these cognitive processes effectively, i.e. applying abstraction, logic, understanding, self-awareness, learning, emotional knowledge, reasoning, planning, creativity, critical thinking, and problem-solving resulting adaptation to new situations, learning from experience, and solving problems.

Artificial intelligence (AI) is a set of technologies enabling computers and machines to learn from the environment, comprehend, perform problem solving, decision making, display creativity and autonomy, seemingly behaving/solving complex tasks as humans do.



Milestones in AI progress

- Early Foundations (1950s):
  - Alan Turing proposed the concept of machine intelligence and introduced the Turing Test (1950).
  - Early AI programs like the "Logic Theorist" focused on problem-solving and symbolic reasoning (FOL).
- Development of Expert Systems (1970s-1980s):
  - Expert systems like DENDRAL, INTERNIST-1 and MYCIN simulated human expertise in specific domains.
  - Inference Engines, Rule-Based Systems, Human-Machine Q&A
  - These systems were used for tasks such as medical diagnosis and chemical analysis.
- Rise of Connectionism (Late 20th Century):
  - Emphasis on artificial neural networks (ANN) to mimic human brain functionality.
  - Advancements in machine learning and pattern recognition.
- Modern AI Advancements (21st Century):
  - Applications include natural language processing, robotics, image recognition, and autonomous systems.
  - Breakthroughs in deep learning and data-driven AI models.

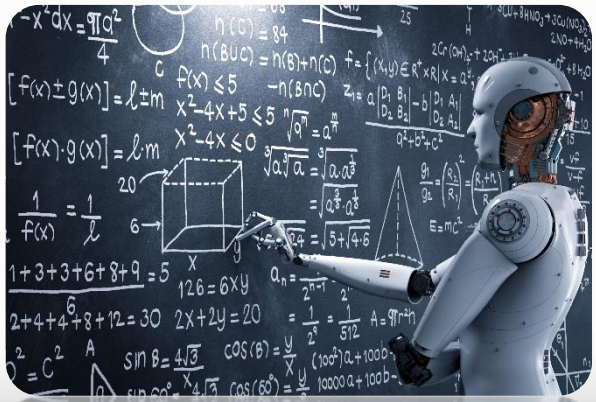


AI Types

Narrow AI (Weak AI) : Type of AI designed and trained to perform a specific task or a narrow range of tasks. It doesn't possess general intelligence and operates within predefined parameters. Examples: Virtual assistants (like Siri or Alexa), recommendation systems, facial recognition software, and self-driving car systems.

***General AI (Strong AI*)**: Also referred as AGI, Hypothetical at this point, representing a machine that has the ability to understand, learn, and apply intelligence across a wide range of tasks, similar to human cognitive abilities. It can adapt and transfer knowledge across domains. Currently, there are no true examples of General AI, though advancements in AI research continue to move in this direction (Open AI O3)

Superintelligent AI: A level of AI that surpasses human intelligence in all areas, including creativity, problem-solving, and decision-making. It could outperform the best human minds in every field.

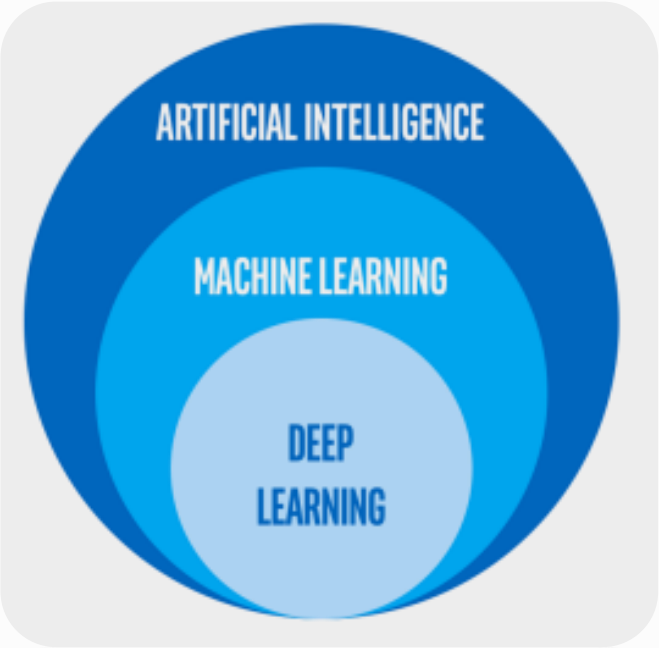


Learning in ML

Learning is ***Improving*** with ***Experience*** at some ***Task***
With respect to some performance Measure

Machine learning is a subset of AI which is incorporated into Devices or software applications - enabling them to be more precise & accurate at finding or predicting outcomes

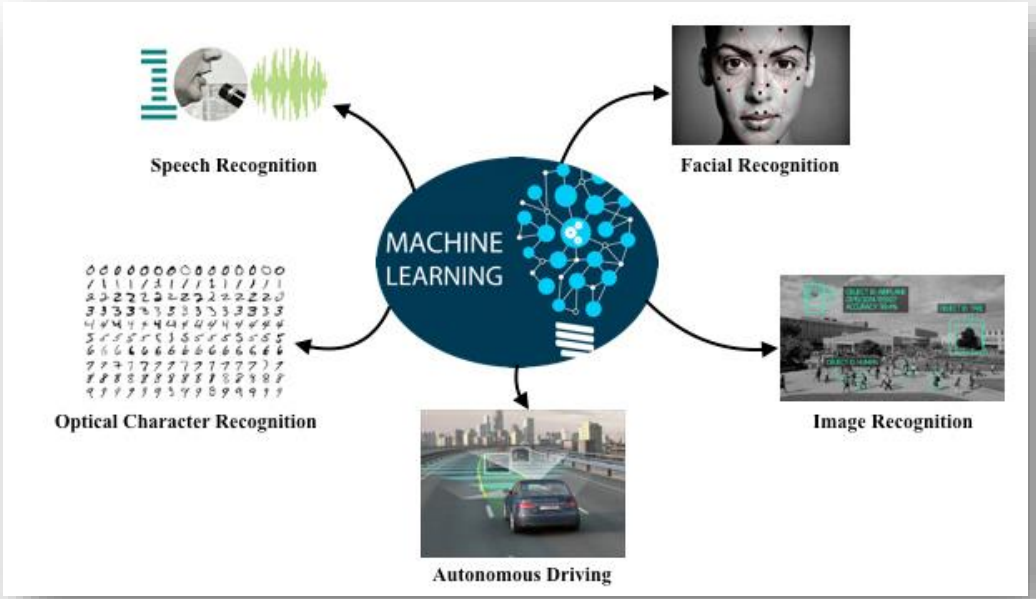
ML used to analyze data and discover patterns, correlations or rules hiding within the data, while AI imitates human cognitive abilities



Machine Learning

"Machine Learning is concerned with computer programs that automatically improve their performance through experience"

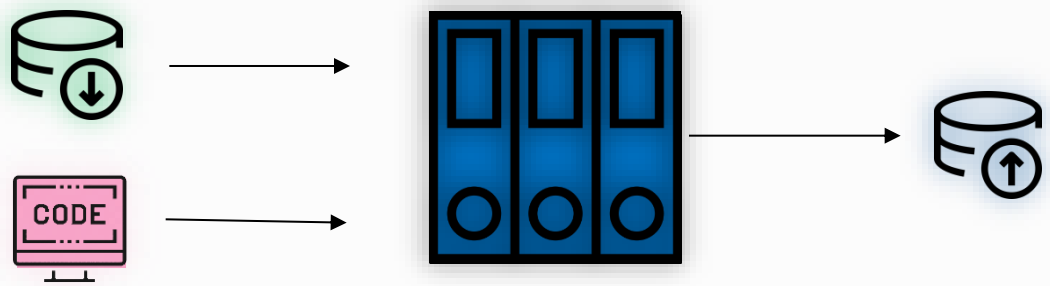
"Learning is any process by which a system improves performance from experience"



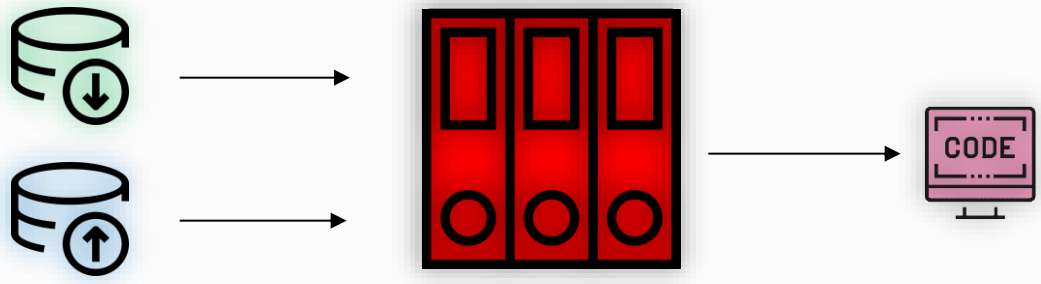
Herbert Alexander Simon
(Nobel Prize Economics 1978)

Machine Learn in a nutshell

20' Centaury (programming)



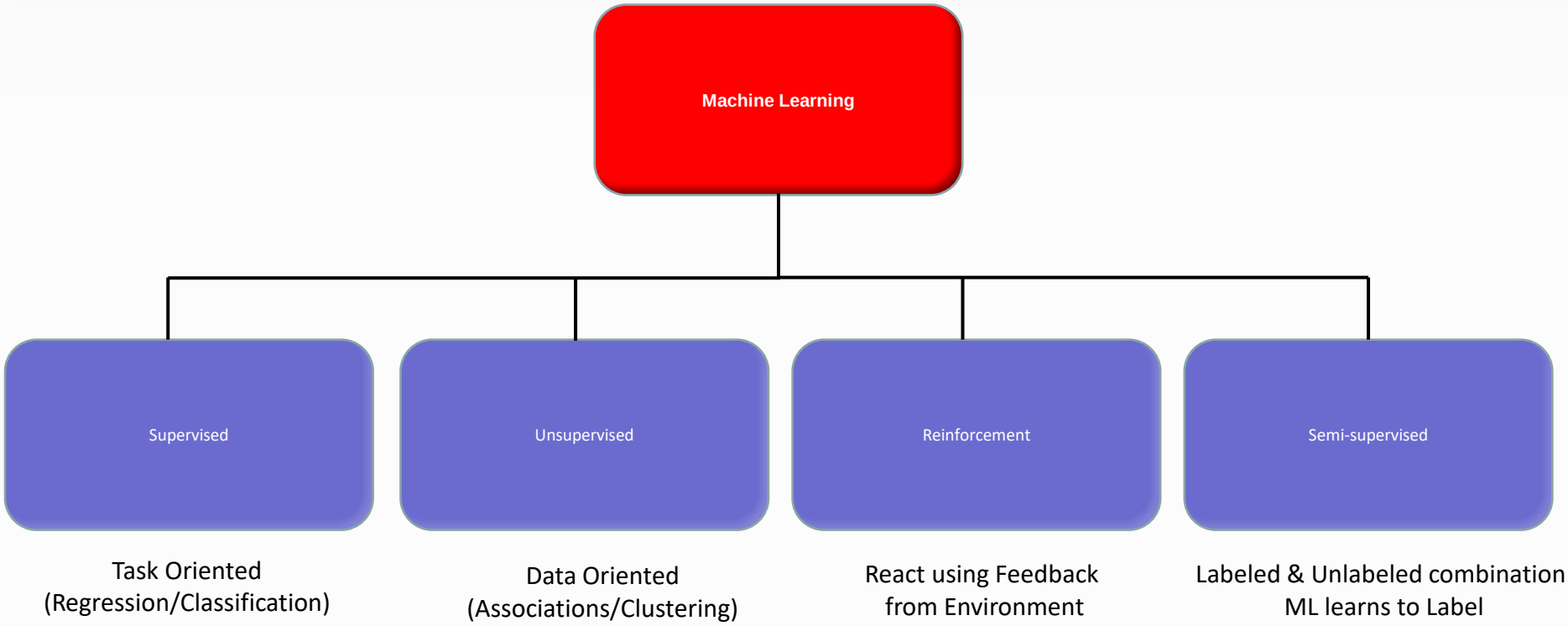
21' Centaury (machine learning)



"For 2019, ML and AI will finally find their way out of the lab and into existing applications one way or another. For the most part, people won't even know it's there because it will be embedded in a seamless fashion." *Scott Parker, Director, Sinequa*.

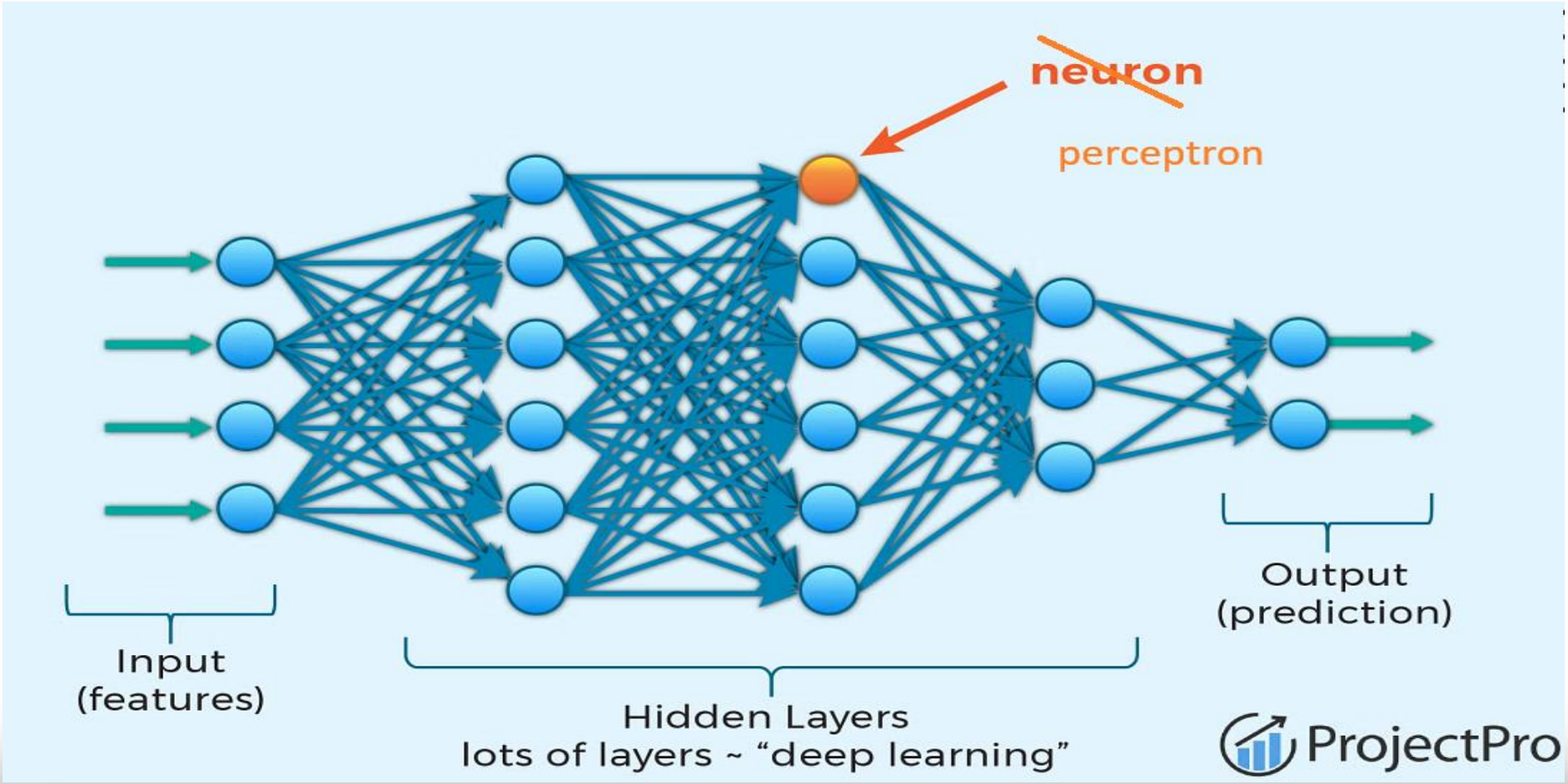
https://dzone.com/articles/aiml-2018-surprises-and-2019-predictions

Types Of Machine Learning



Deep Learning

Deep learning is a subset of machine learning that focuses on utilizing artificial neural networks to perform tasks such as classification, regression, and representation learning.



From early AI to modern AI

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.

MACHINE LEARNING

Machine learning begins to flourish.

DEEP LEARNING

Deep learning breakthroughs drive AI boom.

1950's 1960's 1970's 1980's 1990's 2000's 2010's

Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

Transforms and LLM's

Transformers are a ***type of neural network*** architecture that transforms or changes an input sequence into an output sequence. They do this by learning context and tracking relationships between sequence components

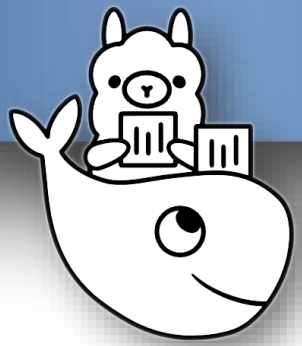
- Transformers excel in ***NLP tasks*** like translation, text generation, and summarization.
- Popular transformer-based models include BERT, GPT, ViTs … But there are many more. Hugging face, for example, holds among 350,000 DL Models, many of them.
- Transformers are the foundation of Large Language Models (LLMs)

https://huggingface.co/models

What is Deep Learning

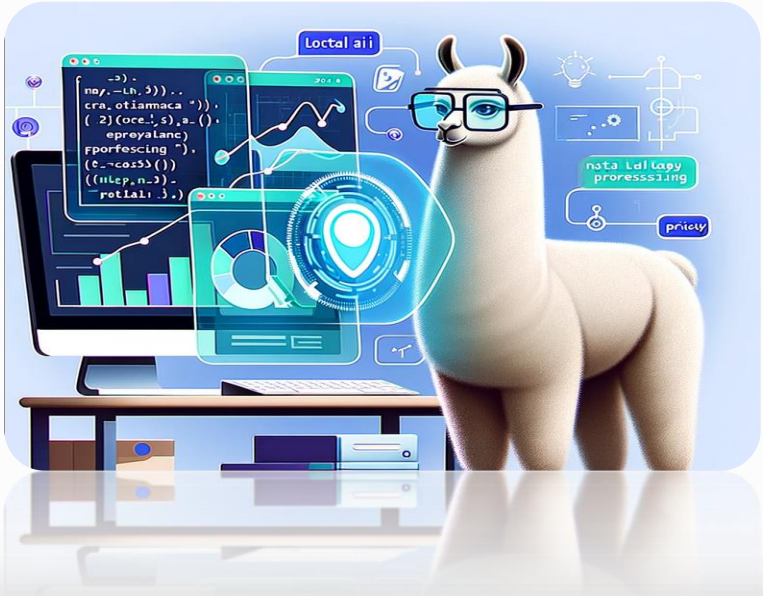
Transformers in Machine Learning - GeeksforGeeks

Ollama

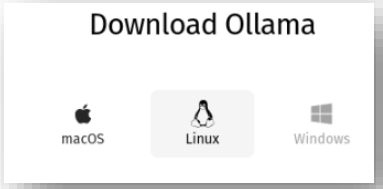


What is Ollama and Docker

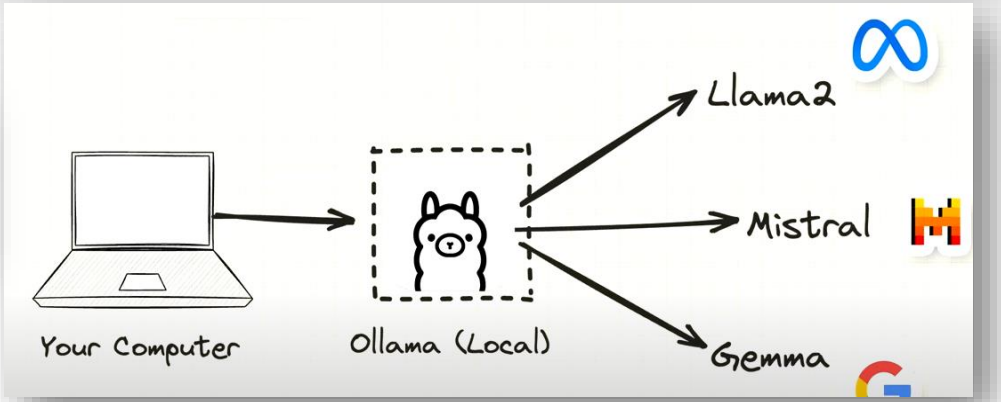
- Ollama enables running **large language models (LLM)** locally
- Ollama can act as your *personal* LLM Advisor
- Local Installation: Ollama Tutorial: Running LLMs Locally Made Super Simple - KDnuggets
- Select LLM Model : library



Local Serving



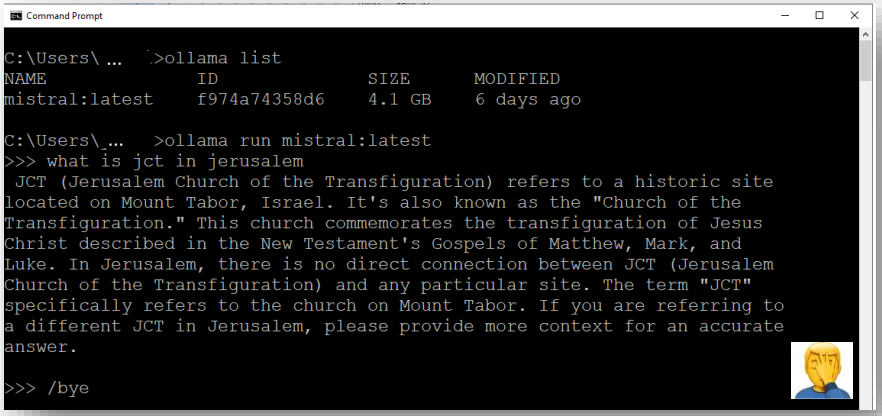

```
ollama pull mistral

ollama pull llama2
```

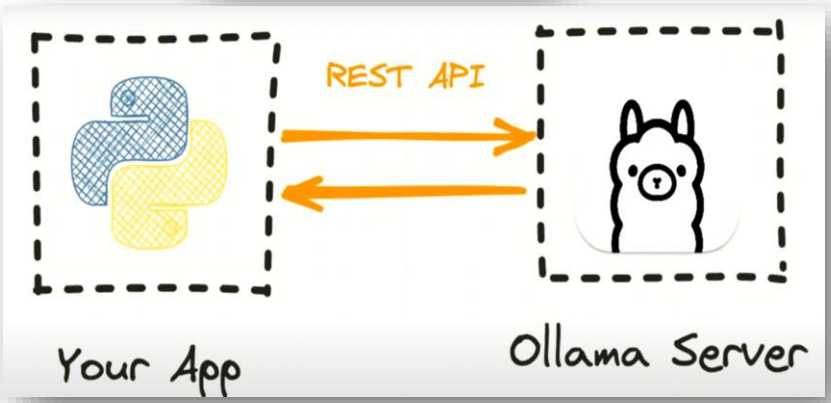


Ollama run mistral

OR

```
ollama serve
```

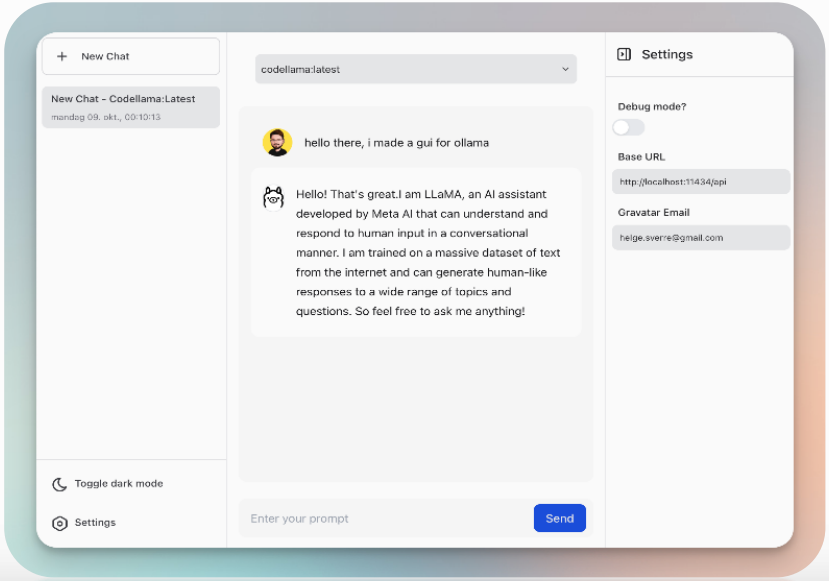




Check: http://localhost: 11434

Ollama

- Navigate to http://localhost:11434 for Ollama's Web GUI
- What can be done ?
  - Send prompts & receive creative responses.
  - Fine-tune your model on your own data.
  - Explore a variety LLM tasks: Translation,question-answering, code completion etc.



Ollama GUI

- Alternativly install open-webui for GUI access:

  open-webui/open-webui: User-friendly AI Interface (Supports Ollama, OpenAI API, ...)

Connect to Ollama using Python

- Create project (folder, virtual environment, activate it)
- pip install ollama

```
import ollama

response = ollama.generate(model='tinyllama',
prompt='what is a model?’)

print(response['response'])
```

RAG with Ollama

RAG Development

- Retrieval Augmented Generation
  - “Cheep” option to make your LLM more specific
- Running Locally
- Database updates
- Validating Output quality
- Potential Tasks:
  - Personal/Enterprise Semantic search
  - Recommendation system
  - Clustering
  - Various NLP tasks
  - ***Planning*** : Integrating AI into Traditional Strategic Planning Models
  - ***Reasoning*** : Reasoning Model (deepseek-reasoner) | DeepSeek API Docs

https://dev.to/mohsin_rashid_13537f11a91/rag-with-ollama-1049

From **automating complex tasks** to **analyzing vast datasets** and **generating intelligent content**, AI agents are leading a tech revolution- and now, you have the chance to be a part of it.

Agentic AI

Agentic AI

Intelligent agents are considered by many to be the ultimate goal of AI. The classic book by Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach (Prentice Hall, 1995), defines the field of AI research as "the study and design of rational agents.

Agentic AI

- ***New*** software solutions ***development paradigm***.
- Uses reasoning & iterative planning to autonomously solve multi-step problems.
- AI chatbots use generative AI to provide responses based on a single interaction, i.e. person enters query and chatbot uses NLP to reply.
- Agentic AI systems ingest vast amounts of data, from multiple data sources and other applications to independently analyze challenges, develop strategies and execute tasks.
- Problem solving process:
  - *Perceive*: Gather and process data from various sources(sensors, databases)
  - *Reason:* LLM acts as reasoning engine – understand tasks, generate solutions, coordinate specific models (content creation, visual processing, recommendation systems).
  - *Act:* Integrate external tools using API's to execute tasks to implement the formulated plan.
  - *Learn:* continuously improves through a feedback loop. data generated from its interactions is fed into the system to enhance models

Agentic AI Examples

- ***Customer Service:*** Improve customer support by enhancing self-service capabilities and automating routine communications.
- ***Content Creation:*** Quickly create high-quality, personalized marketing content.
- ***Software Engineering:*** Boosting developer productivity by automating repetitive coding tasks Or even blueprinting a full system.
- ***Healthcare:*** Distill critical information from vast amounts of patient data, to help physicians make better-informed care decisions.
- ***Media analytics***: Boost capabilities of workforces relying on visual information from devices (cameras, IoT sensors and vehicles**).**
- **Scientific Bigdata Analytics:**