# Teaser

ChatGPT

DreamStudio

# Most Famous AI/ML Applications

recommendations

chatbots

ChatGPT Bard

autonomous driving

WAYMO TESLA

translation

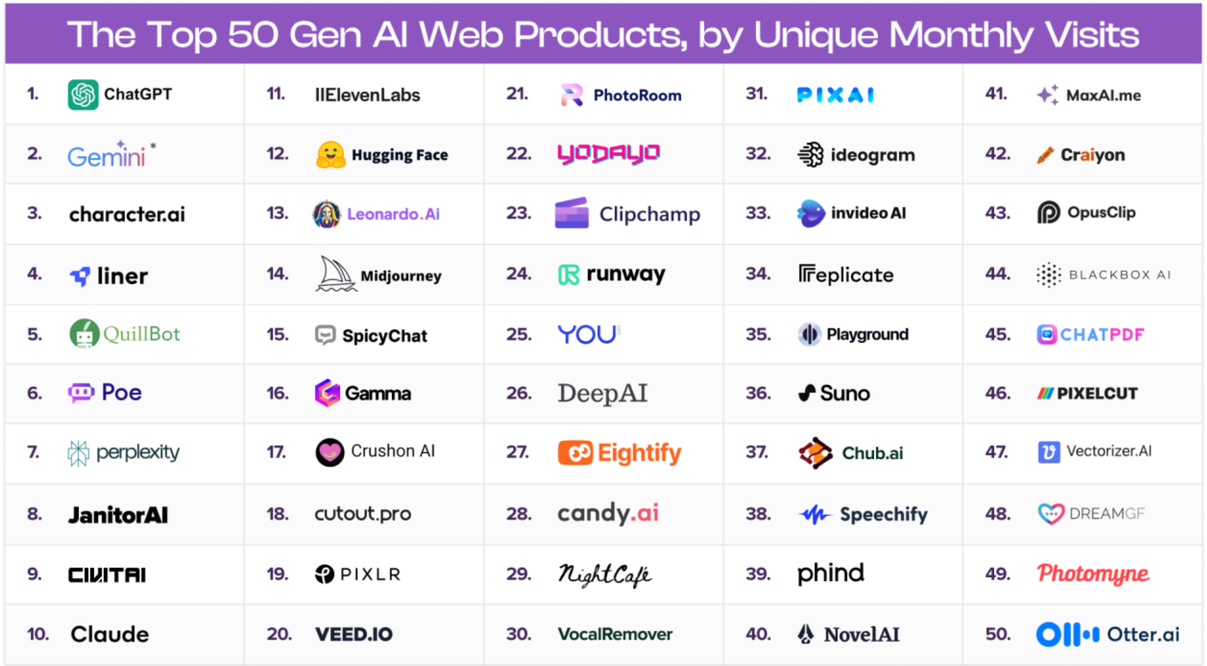
DeepL

robotics

OCR

assistants (speech recognition)

and many more …



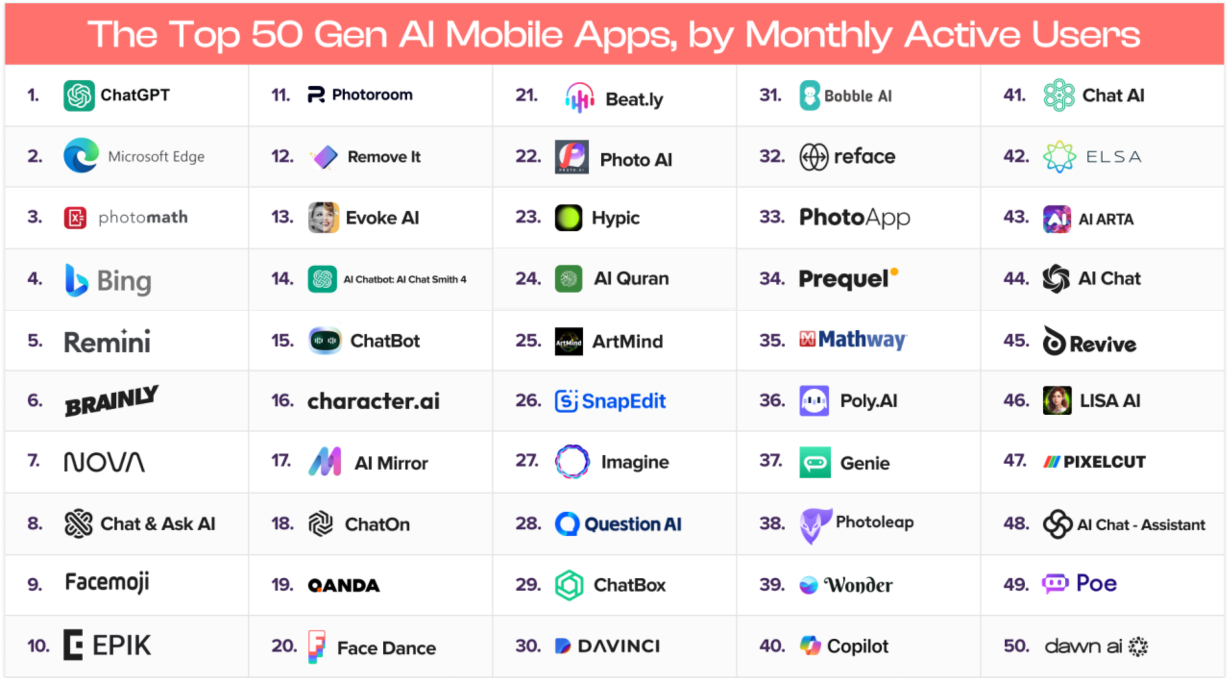
## The Top 50 Gen AI Web Products, by Unique Monthly Visits

| | | | | |
|---|---|---|---|---|
| 1. ChatGPT | 11. ElevenLabs | 21. PhotoRoom | 31. PIXAI | 41. MaxAI.me |
| 2. Gemini* | 12. Hugging Face | 22. yodayo | 32. ideogram | 42. Craiyon |
| 3. character.ai | 13. Leonardo.Ai | 23. Clipchamp | 33. invideo AI | 43. OpusClip |
| 4. liner | 14. Midjourney | 24. runway | 34. replicate | 44. BLACKBOX AI |
| 5. QuillBot | 15. SpicyChat | 25. YOU | 35. Playground | 45. CHATPDF |
| 6. Poe | 16. Gamma | 26. DeepAI | 36. Suno | 46. PIXELCUT |
| 7. perplexity | 17. Crushon AI | 27. Eightify | 37. Chub.ai | 47. Vectorizer.AI |
| 8. JanitorAI | 18. cutout.pro | 28. candy.ai | 38. Speechify | 48. DREAMGF |
| 9. CIVITAI | 19. PIXLR | 29. NightCafe | 39. phind | 49. Photomyne |
| 10. Claude | 20. VEED.IO | 30. VocalRemover | 40. NovelAI | 50. Otter.ai |

*formerly Bard

Charts are for informational purposes only and should not be used for investment decisions. Past performance is not indicative of future results. None of the above should be taken as investment advice; see a16z.com/disclosures.

a16z Consumer



## The Top 50 Gen AI Mobile Apps, by Monthly Active Users

| | | | | |
|---|---|---|---|---|
| 1. ChatGPT | 11. Photoroom | 21. Beat.ly | 31. Bobble AI | 41. Chat AI |
| 2. Microsoft Edge | 12. Remove It | 22. Photo AI | 32. reface | 42. ELSA |
| 3. photomath | 13. Evoke AI | 23. Hypic | 33. PhotoApp | 43. AI ARTA |
| 4. Bing | 14. AI Chatbot: AI Chat Smith 4 | 24. AI Quran | 34. Prequel | 44. AI Chat |
| 5. Remini | 15. ChatBot | 25. ArtMind | 35. Mathway | 45. Revive |
| 6. BRAINLY | 16. character.ai | 26. SnapEdit | 36. Poly.AI | 46. LISA AI |
| 7. NOVA | 17. AI Mirror | 27. Imagine | 37. Genie | 47. PIXELCUT |
| 8. Chat & Ask AI | 18. ChatOn | 28. Question AI | 38. Photoleap | 48. AI Chat - Assistant |
| 9. Facemoji | 19. QANDA | 29. ChatBox | 39. Wonder | 49. Poe |
| 10. EPIK | 20. Face Dance | 30. DAVINCI | 40. Copilot | 50. dawn ai |

Charts are for informational purposes only and should not be used for investment decisions. Past performance is not indicative of future results. None of the above should be taken as investment advice; see a16z.com/disclosures.

a16z Consumer

source

# Literature

If you want to go a bit deeper …

- Deep Learning

- The Little Book of Deep Learning

- Understanding Deep Learning

# Introduction AI/ML

# Main Areas of Artificial Intelligence

from wikipedia

- **computer vision**

data: spatial structures (e.g., images), SOTA: Convolutional Neural Networks (CNN)

- **natural language processing**

data: sequential structures (e.g., text), SOTA: transformers

- **automated decision making, robotics**

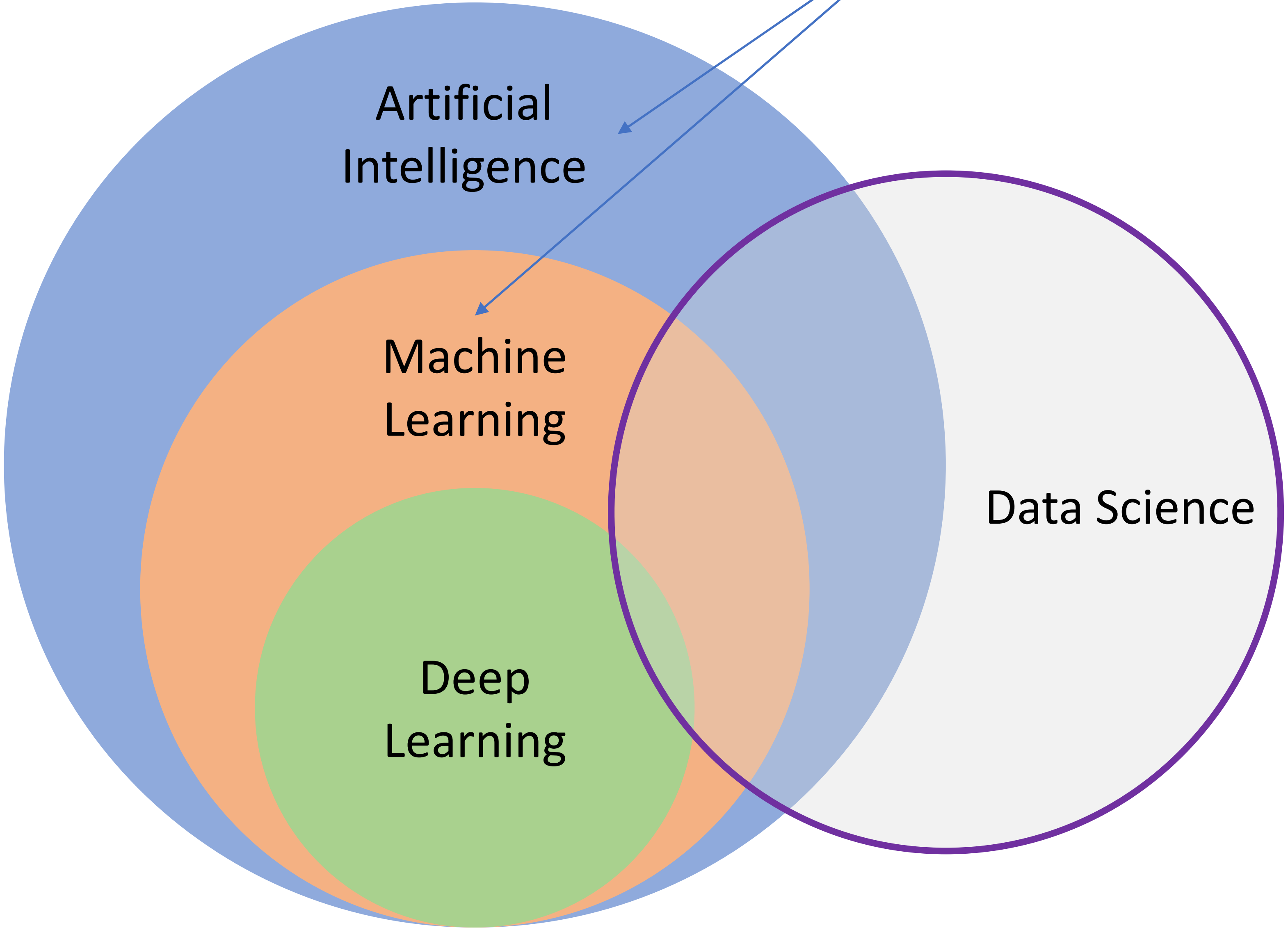
data: sequential actions (e.g., games), SOTA: reinforcement learning

agency:
perception – thought – action

All of these are enabled by one key ingredient:

*learning from experience/data* (**Machine Learning**)

data can also be tabular (structured): columns as features, rows as independent samples blend of diverse components from different domains (statistics, optimization, computer science, …)



*Deep Learning*:

special kind of ML algorithms using *deep* neural networks (e.g., CNNs, transformers)

*Data Science*:

extract knowledge from data (by means of ML, among other things)

# ML: Learning from Experience/Data

mainly exploiting statistical dependencies with the aim of **generalization** to new (e.g., future) data (compare with human reasoning by analogies)

training (usually offline optimization):

**ML algorithm + data = explicit algorithm** (to be used at inference time)

→ reduction of complexity and much better generalizability compared to handcrafted algorithms

analogy: Humans do not hit the ground running (storage capacity of DNA limited) but have learning capabilities.

# Ladder of Generalization

shallow learning:

representation encoded in features

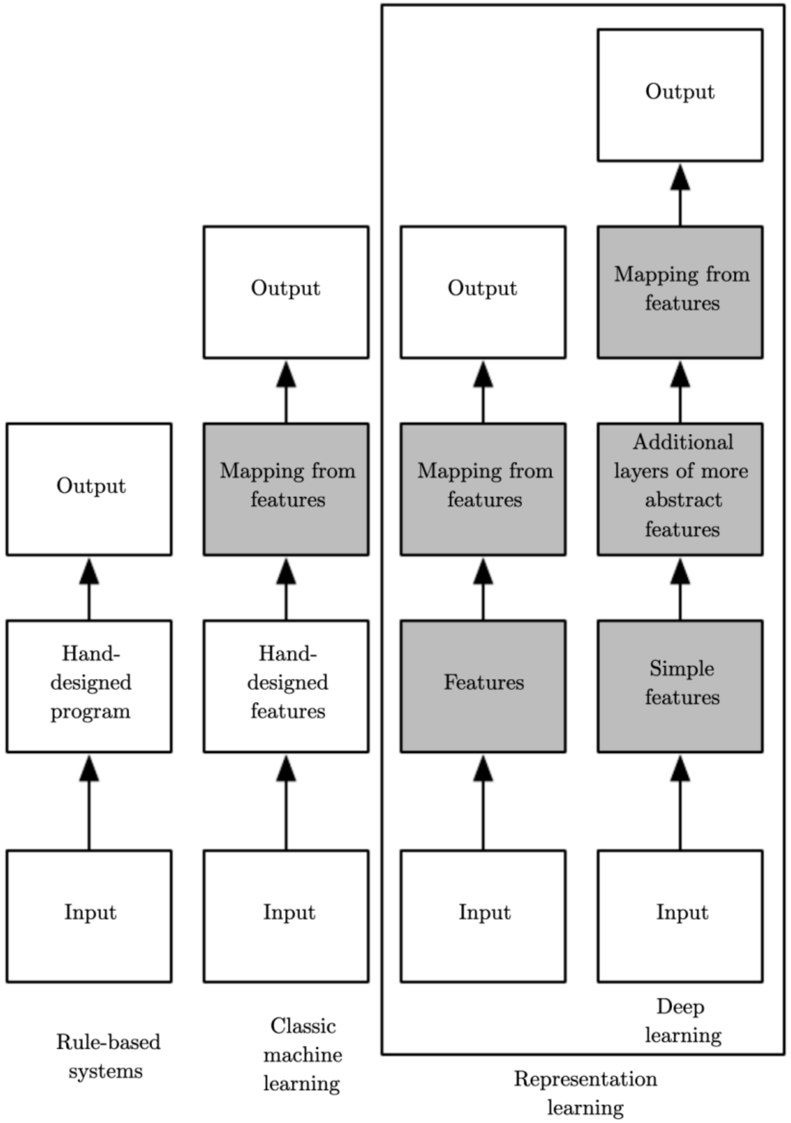
→ feature engineering

deep learning:

representation encoded in network

→ feature/representation learning

(hierarchy of concepts learned from raw data in deep graph with many layers)

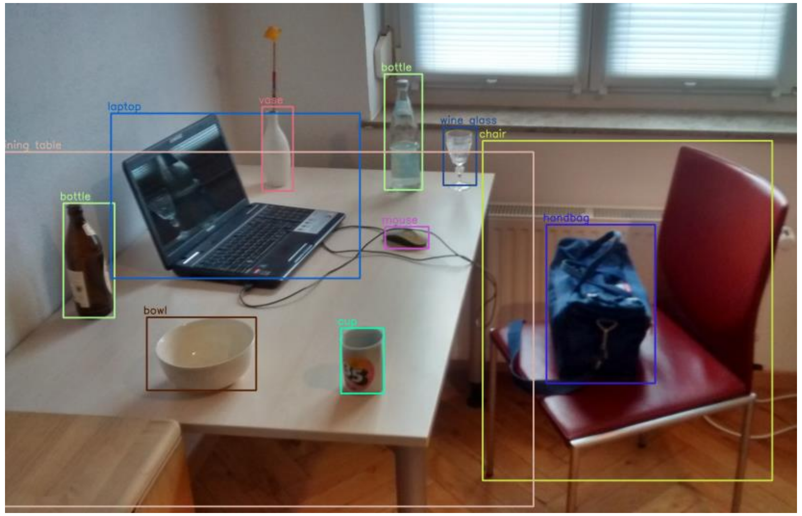


source

# When to Use ML (= Learning from Data)

## automation

too complex for rules



from wikipedia

object recognition, chat bot, …

## complexity / uncertainty

too complex for humans

AlphaFold

protein structure predictions, demand forecasting, …

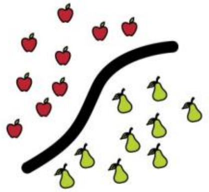
more scientific use cases: medicine (imaging, diagnosis, drug design), particle physics (analysis of collider experiments), material science (material properties and design of new materials), …

# MACHINE LEARNING

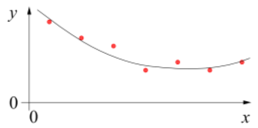
*training target available (labeled or past data)*

## SUPERVISED

### CLASSIFICATION



### REGRESSION



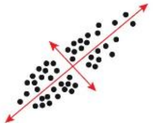
*data not labeled in any way*

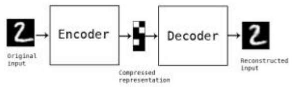
## UNSUPERVISED

### CLUSTERING

### GENERATIVE MODELS

### DIMENSIONALITY REDUCTION





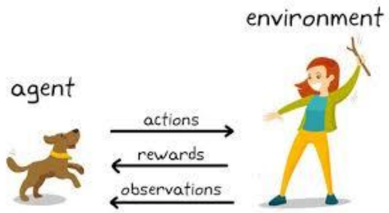
### ASSOCIATION

*no supervision, but goal-based interaction with environment*

## REINFORCEMENT LEARNING

### LEARN STATE OR ACTION VALUES

### LEARN POLICY DIRECTLY



learning by teacher
(high-dimensional curve fitting)

learning by observation
(pattern recognition)

learning by trial-and-error
(sequential decision making)

unsupervised and reinforcement learning can both be cast as supervised-learning setup

# Supervised Learning

**learning by teacher** → usually rather narrow tasks (passive approach)

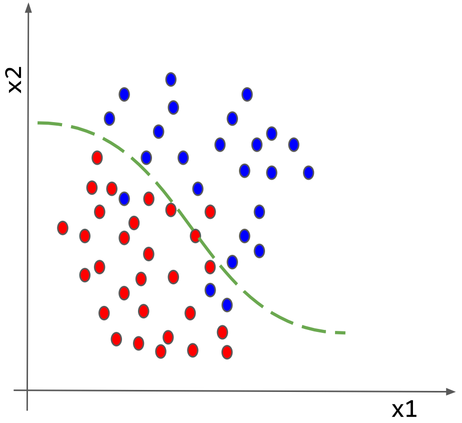
### Target Quantity

- **known in training**: labeled samples or observations from past
- to be **predicted** for unknown cases (e.g., future values)

### Features

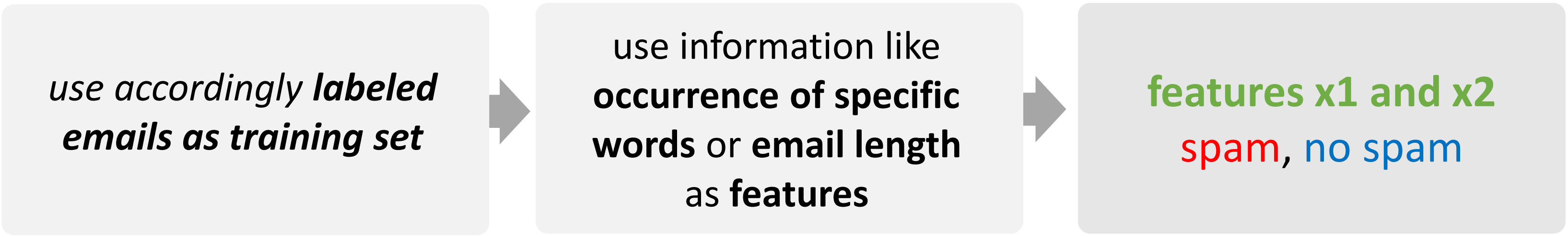
input information that is

- correlated to target quantity
- known at prediction time



### Example: Spam Filtering

*Classify emails as spam or no spam*



*use accordingly **labeled emails as training set***

use information like **occurrence of specific words** or **email length** as **features**

**features x1 and x2**
spam, no spam

# But Before: Data Processing

environment:

WSL, Python, virtualenv, pip

scientific Python stack:

NumPy, pandas, matplotlib

coding example: stock market data