NOVENTIQ
Global expertise, local outcomes

AI Is Here – The Journey from Challenges to Secure Business Growth

Bozó Károly Barnabás

2025. Nov 4.

We live in the Era of AI

***Whatever we do* – we run into AI:**

Social media, Search engines, smart cars, smart fridges, chess programs, chat bots, smart lightbulbs, LLM, SLM

Companies build solutions to leverage the benefits of Artificial Intelligence:

- *Process/ task automation*
- *Decision making*
- *Analytics*
- *Predictions*
- *Content generation*

Agenda

What is AI

- Hello world!- neural network
- Basics of Language models

Demo

- RAG Solution

How to drive succesfull Ai implementations

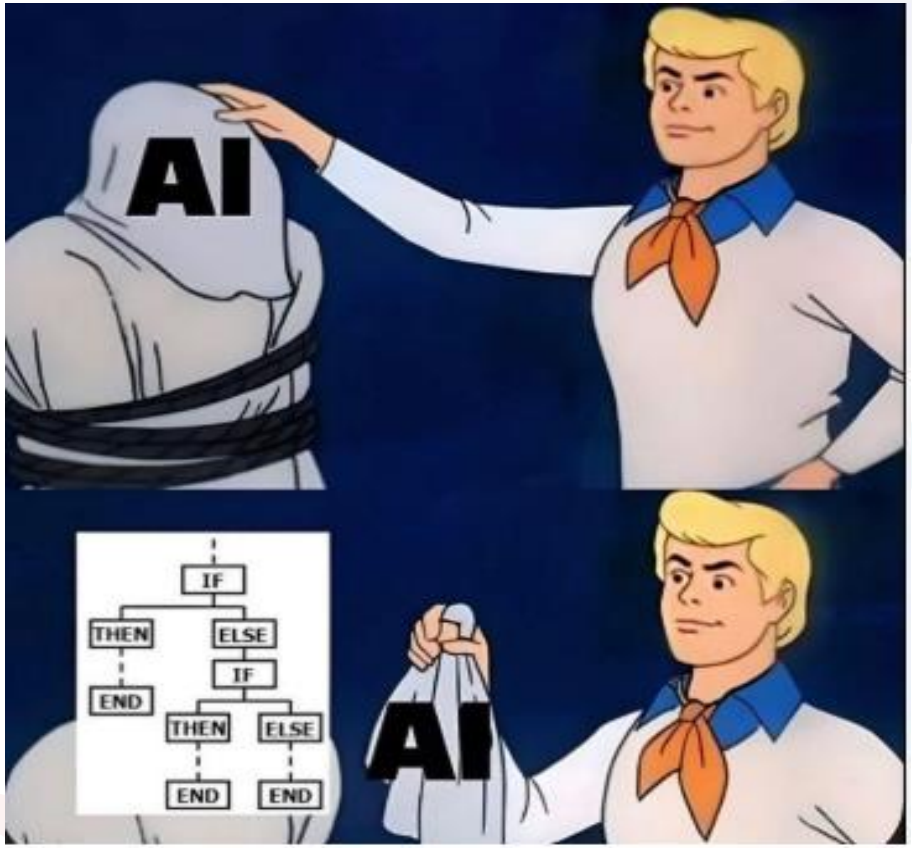
- Industry best practices

What makes Ai so different

Conceptual differences between programming and Neural networks

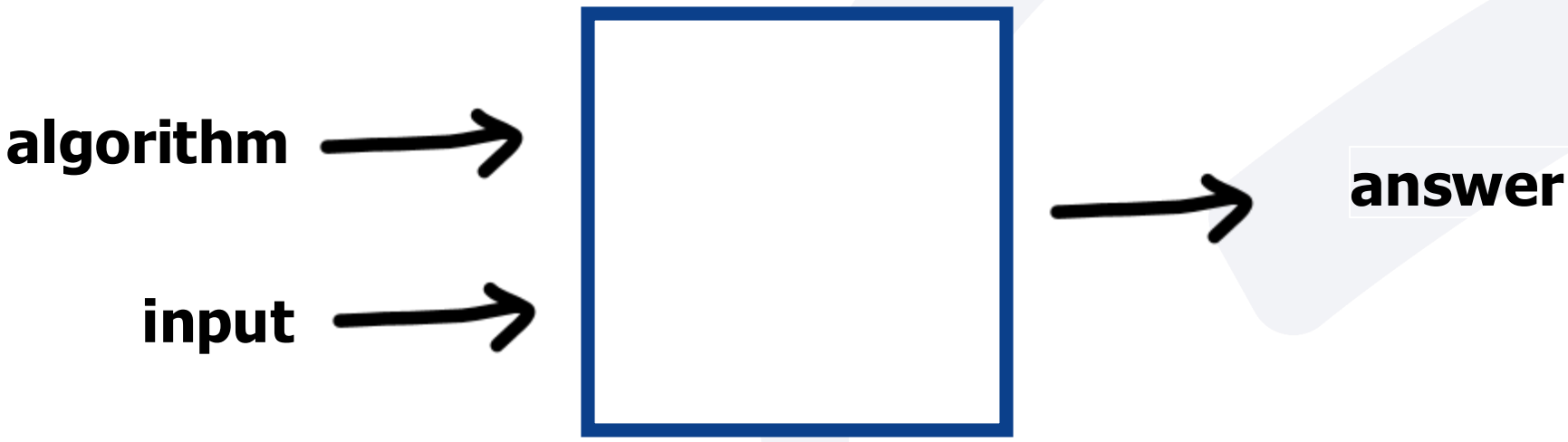
Some of them are different – they are not all smart the same way:

- *Social media,*
- *Search engines,*
- *Smart cars,*
- *Smart fridges,*
- *Chess programs,*
- *Chat bots,*
- *Smart lightbulbs,*
- *LLM,*
- *SMLS…*



Programing

We write the algorithm

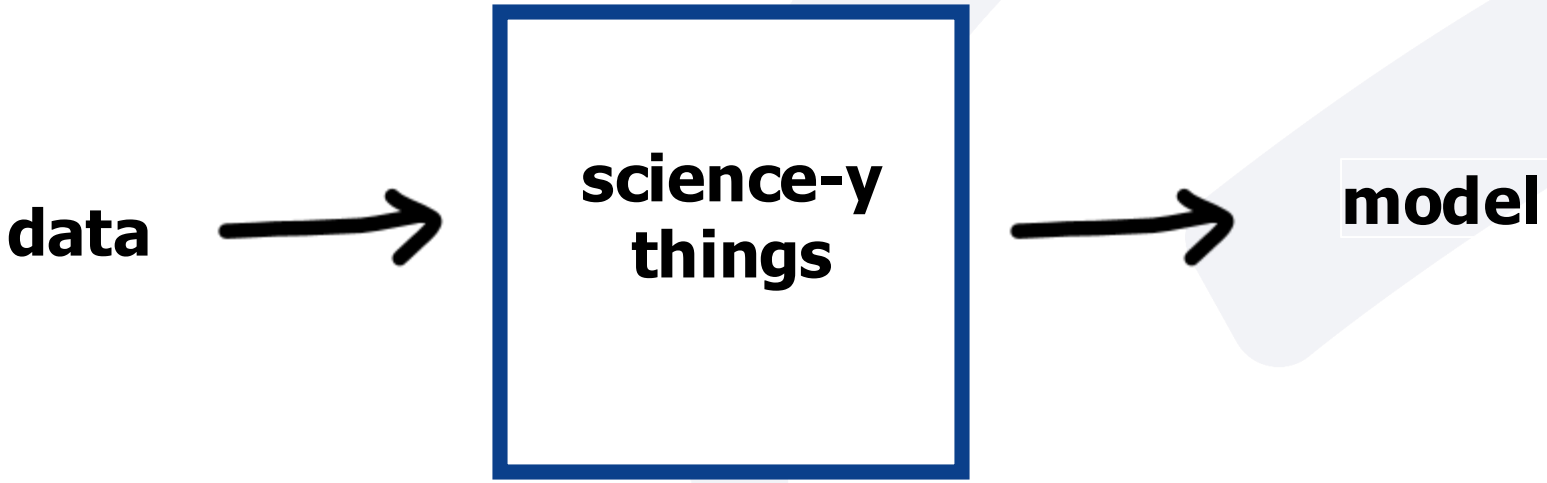


Neural Networks

We do not write the algorithm

Neural Networks

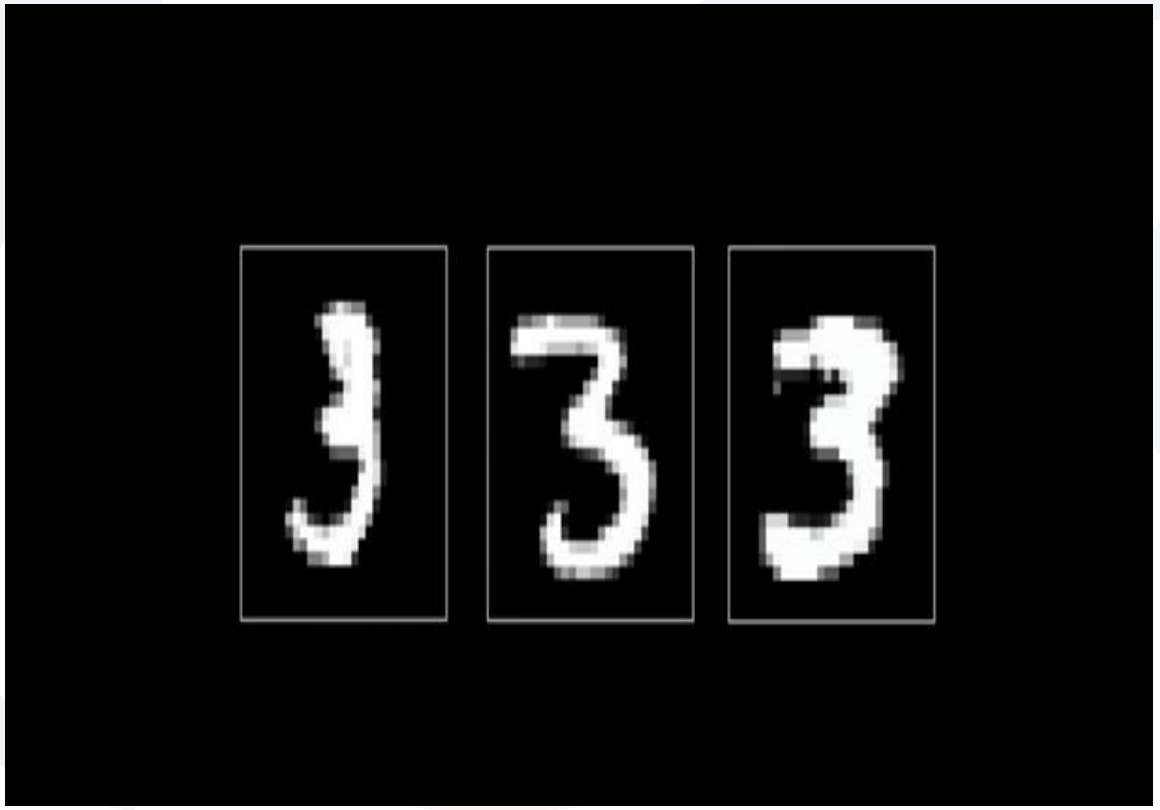
We provide data and what happens in the box is unknown



Hello World – for neural networks

The aim is to have an algorithm that can recognize handwritten digits from 0-9

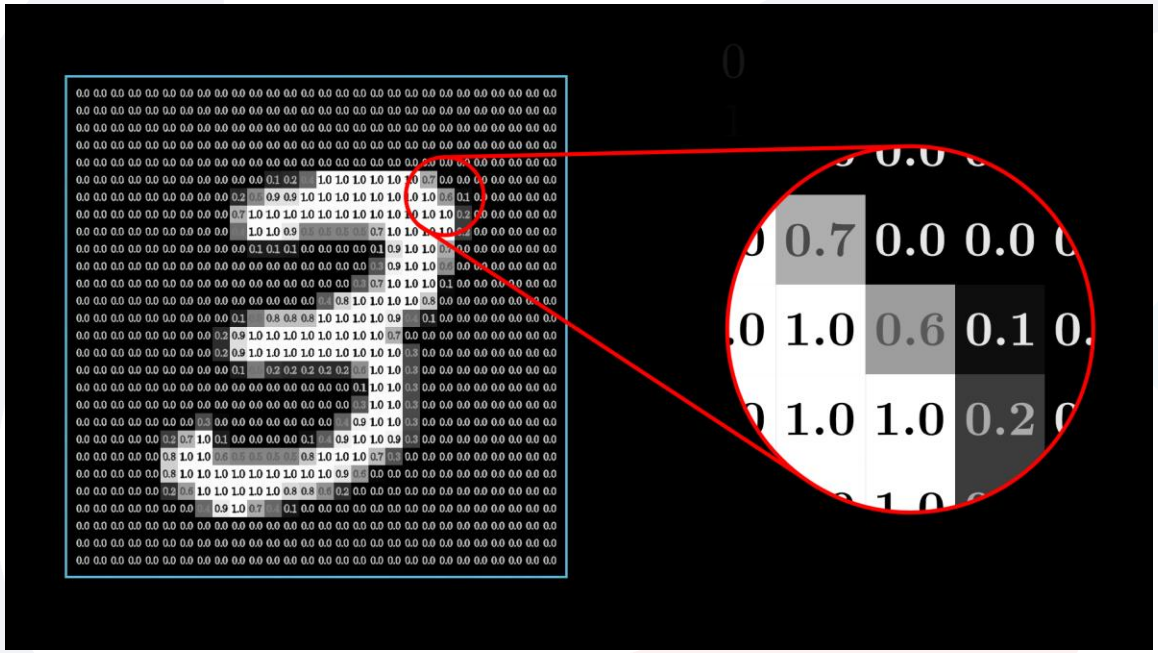
- Seems like a simple task
- The human brain can easily identify all these numbers as 3s
- The traditional methods of computer programming, (if statements, for loops, classes, objects and functions,) just don't seem suitable to tackle this problem



The Numbers

28x28 pixel to start with

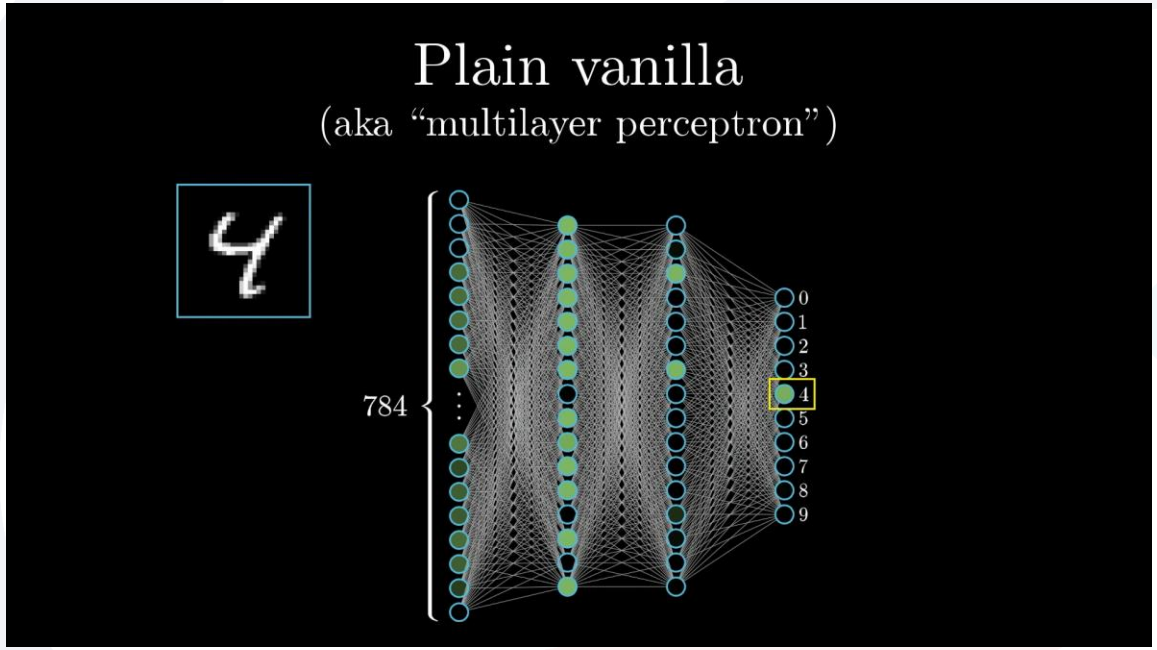
- In our case every picture consists of 28x28 (784) pixels
- Every pixel holds a number between 0 and 255 on a gray scale
- We divided these numbers by 255



Neural Networks

It all comes down to vector calculations in reality

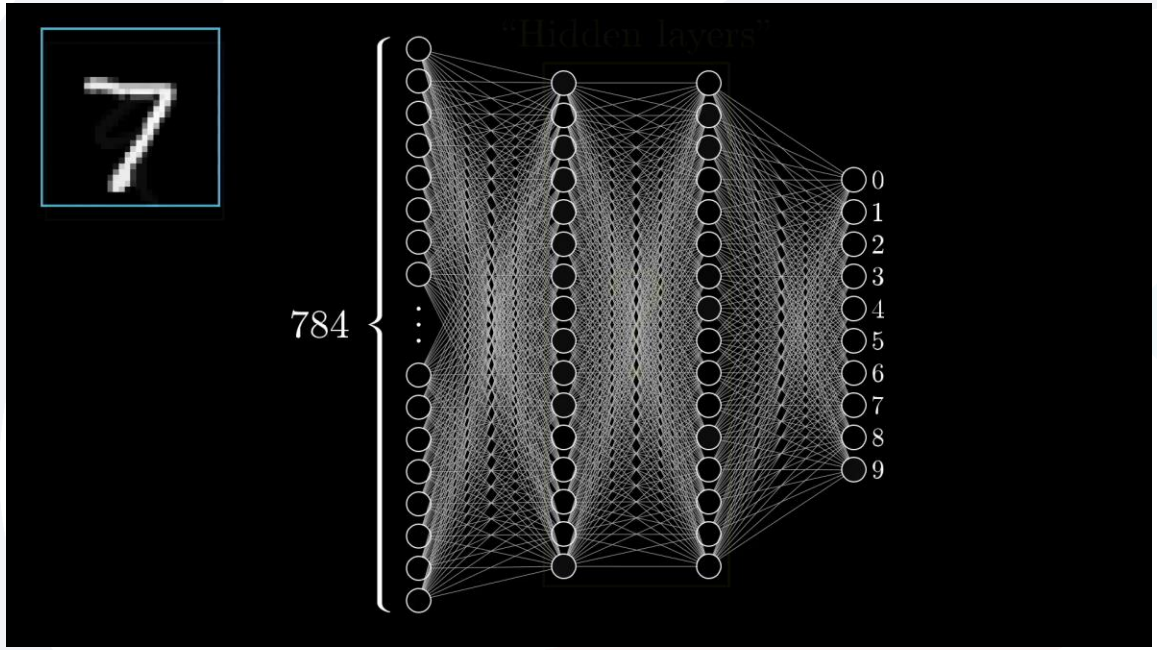
- Input layer: 28x28=784
- Hidden layers: 2x16
- Output layer: 10



How the Network work

The neuron activation run from left to right

- Every neuron is connected to all neurons in the previous layer
- The activation of one neuron contributes to the activation of the neurons in the other layers
- Overall, the activation of one layer determines the activation of the next layers



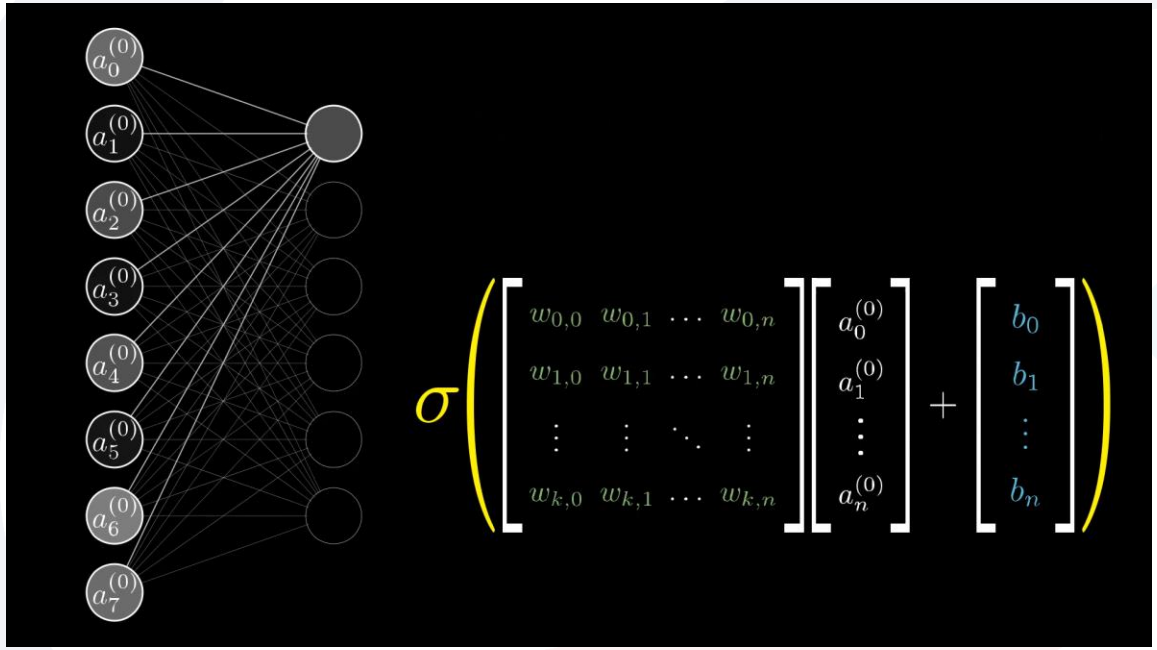
Weighted sums + Biases

Vector calculations

- 3 ingredients:
  - Activation number
  - Weighted sum of the connections
  - Bias- random number to alter activation

- Wrapped around the Sigmoid function

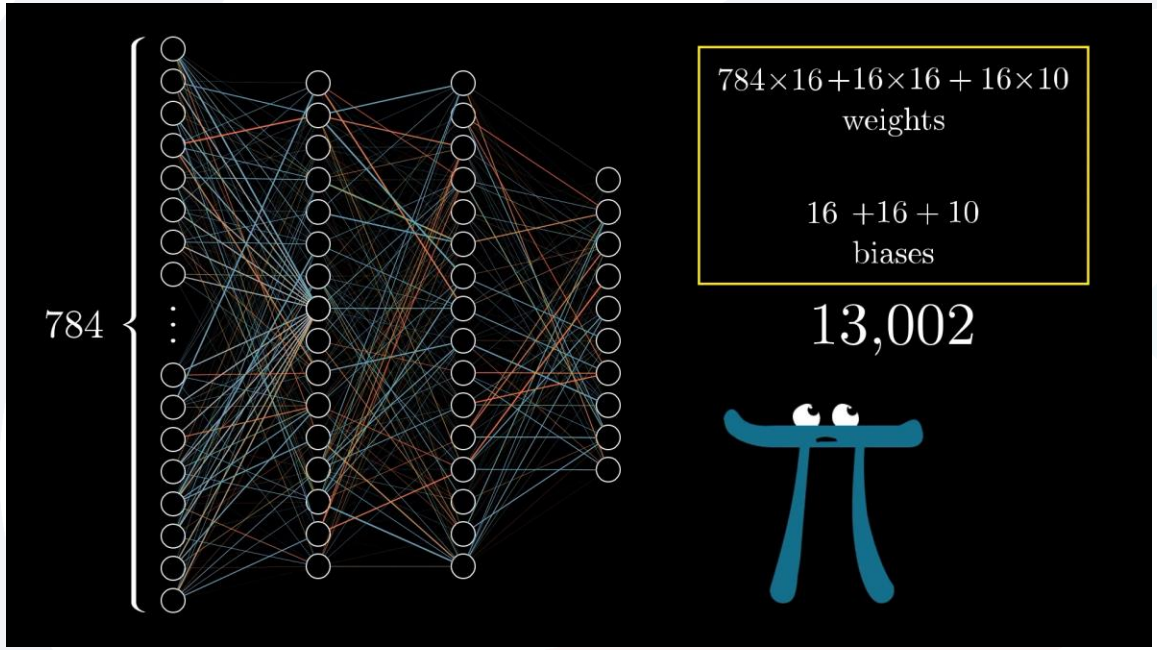
- All organized as vectors and columns for efficiency



The final Model

Layers and parameters

1 input layer
2 hidden layers
1 output layer
+
13002 paremeters

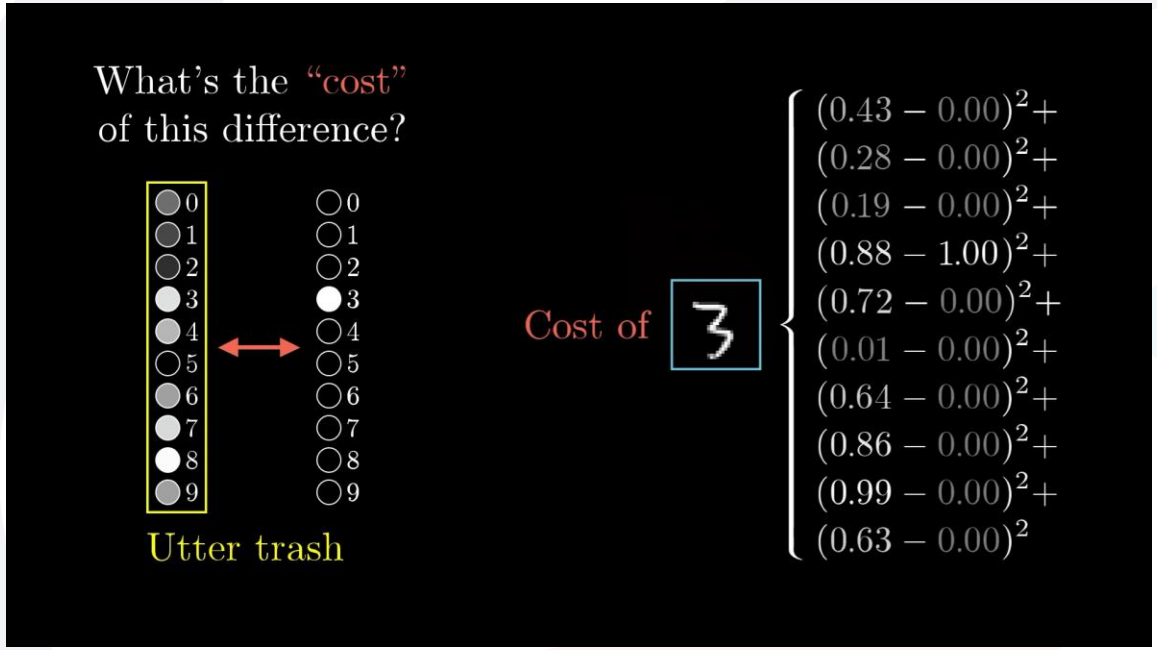


The cost of the model

The sum of the squares of the differences

- The msg:
  - "No, bad computer, that output layer should have activations which are 0 for most neurons, but with a 1 for the third neuron. What you gave me is utter trash!"
- Mathematically:
  - We add up the squares of the differences between each of these output activations and the values you want them to have. We'll call this the "cost" of a single training example.

- The cost is small when the network confidently classifies this image correctly but large when it doesn't know what it's doing.

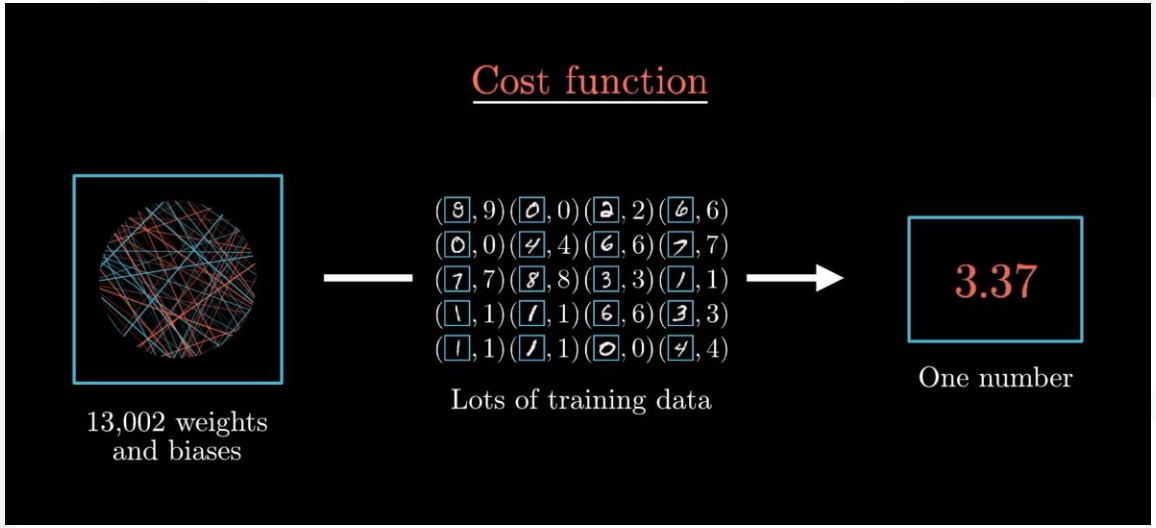


The cost function

We add up these invidual costs

Cost function

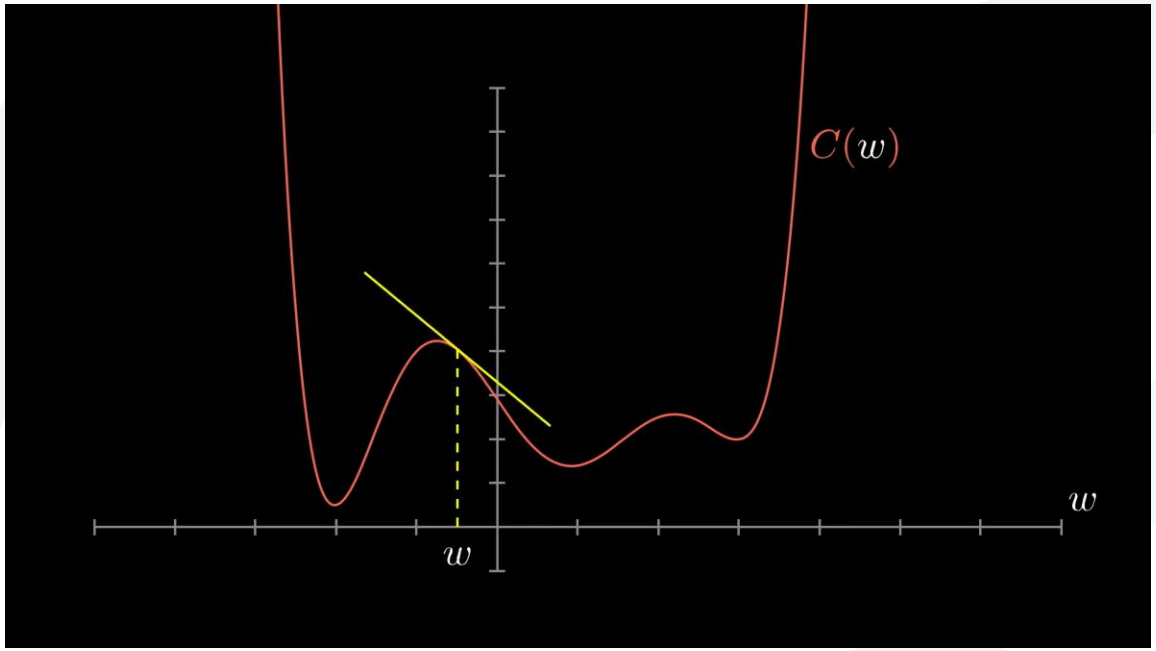
- Input: 13002 weights and biases
- Output: 1 number
- Parameters: All the training examples



The minimum of a function

Simple example in the 2 dimensons space

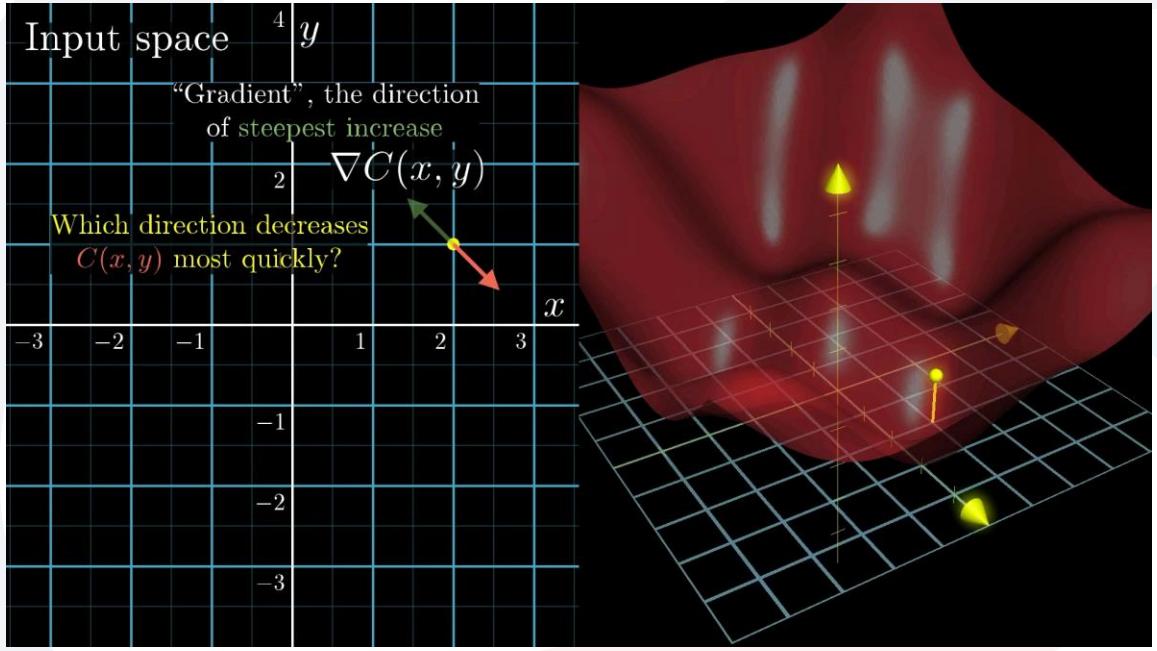
- There is 1 number as an input and 1 number as an output
- The actual calculation comes down to derivatives
- Finding the slope of the tangent
- Where the slope of the tangent line is 0 that is the minimum
- The task is to find the minimum and overcome the local minimums



Gradient descent with 13002 inputs

Gradient nudge weights

- We are looking for the gradient descent of the cost function
- To optime the cost function we have to do vector calculations in the 1302 vector space
- The minimum of the cost function will provide the best results of the models functionality



Neural Networks and Deep Learning

The content of these slides were based on:

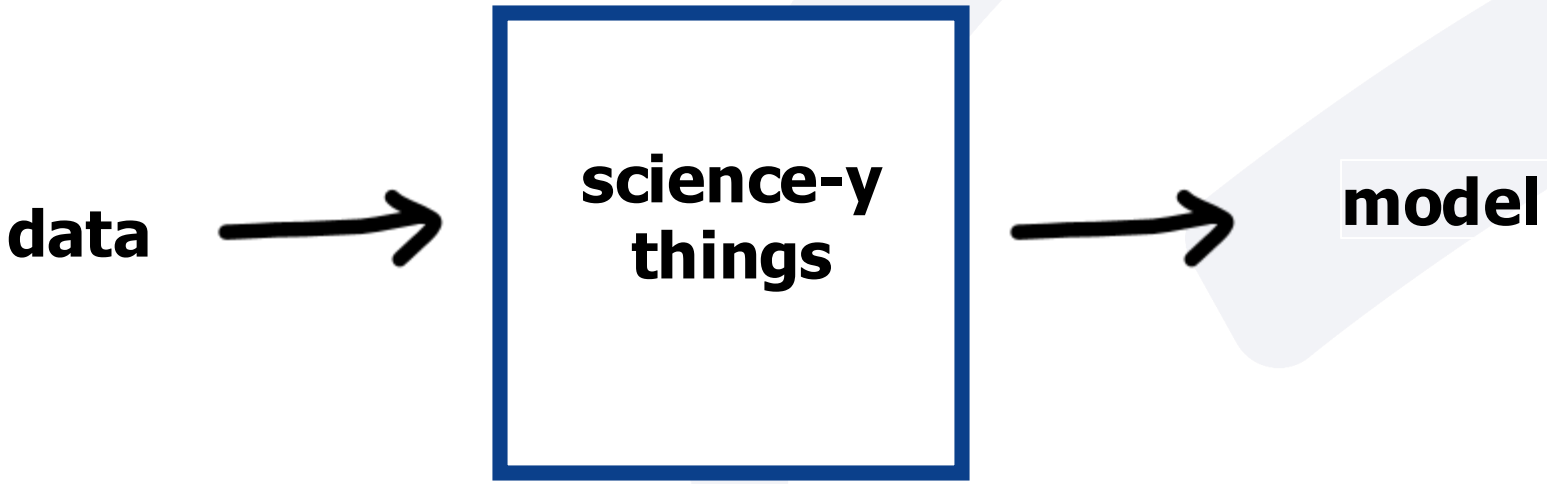
«Neural Networks and Deep Learning»
written by Michael Nielsen

Gifs and pictures were prepared by:

3Blue1Brown

Deep learning

We provide data and what happens in the box is unknown



Simple MNIST Net vs ChatGPT 3.0 (GPT-3)

When size does matter – A quick scale comparison

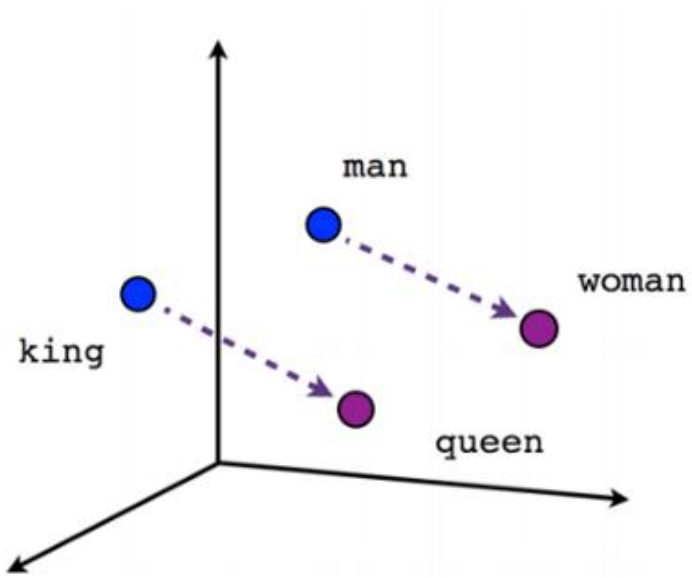
Category	Simple MNIST digit net	ChatGPT 3.0 (GPT-3)
Data needed to train	MNIST: 60,000 28×28 grayscale images (+10,000 test) ≈ 47 MB raw pixels	≈ 45 TB
Number of parameters	• 13002	175 B
Model size (weights for inference)	≈ 0.25–7 MB (FP32) Smaller with FP16/INT8	Very large: ≈ 700 GB (FP32) ≈ 350 GB (FP16) ≈ 175 GB (INT8) (Served sharded across many GPUs/TPUs)

Actions speak louder than words...

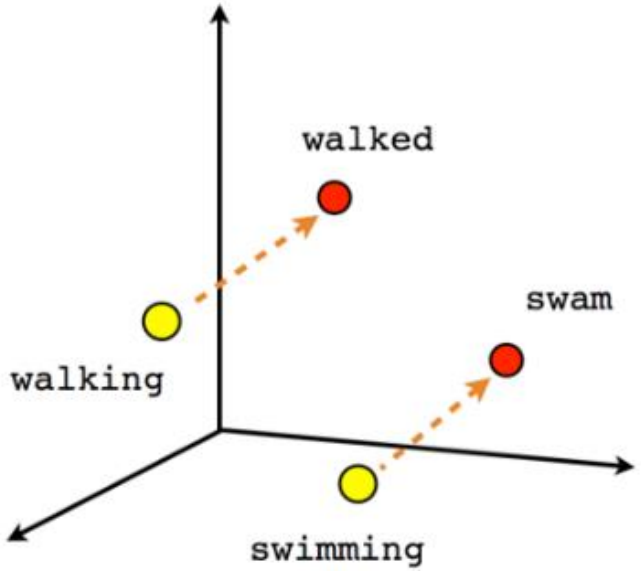
All roads lead to Rome... _ _ _ _

Vector Embedding examples

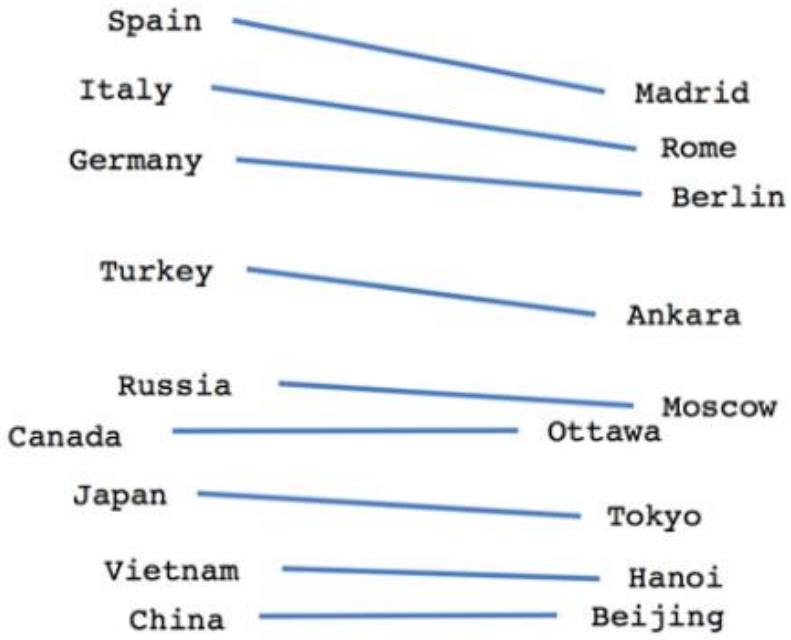
In the 3 dimensonal space



Male-Female



Verb tense



Country-Capital

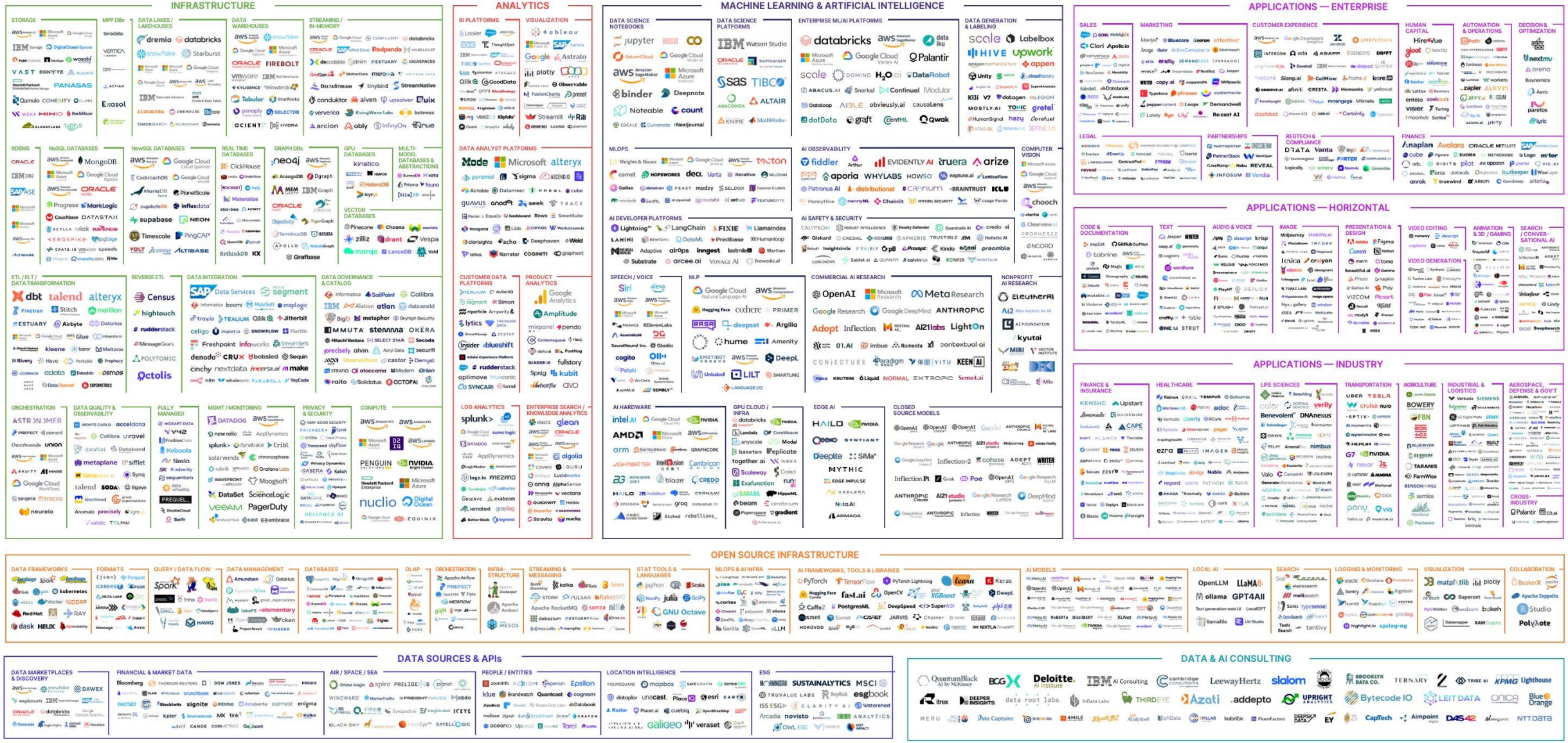
[-0.013595974072813988, -0.02448209747672081, -0.024014882743358612, -0.005469342693686485, 0.009309271350502968, 0.017380420118570328, -0.04211948812007904, -0.025509972125291824, -0.014273436740040779, 0.03728380426764488, 0.0036121606826782227, -0.029154254123568535, -0.01878206804394722, -0.011224854737520218, 0.003352272091433406, -0.0017345380038022995, -0.040554314851760864, 0.0009862629231065512, -
0.017754193395376205, -0.001733077922835946, 0.051907654851675034, -0.002927397843450308, 0.00038581789704039693, 0.028570234775543213, 0.00978232640773058, 0.005218214355409145, -0.01813964545726776, -0.007831702008843422, -0.0231154914945364, -0.0025463250931352377, -0.01759066805243492, 0.005790553521364927, 0.0014286576770246029, -0.02476242743432518, 0.010699237696826458, 0.02707514539361, -0.03326575085520744,
0.009869929403066635, -0.01954129338264653, 0.006943991873413324, 0.05260847881436348, -0.011551906354725361, -0.006856388878077269, -0.023851357400417328, -0.003282189602032304, 0.055365048348903656, -0.039362914860248566, -0.0009563319617882371, -0.000542408146429807, 0.013245562091469765, -0.001679056091234088, -0.0008862495888024569, 0.024645622819662094, 0.0023915600031614304, -0.011143092066049576,
0.008614287711679935, 0.019167520105838776, 0.005898597184568644, -0.03779774159193039, -0.03506452962756157, -0.026911618188023567, 0.025813661515712738, 0.022473070770500209, -0.0074170478619635105, -0.006704544182866812, -0.006389173679053783, 0.008882937021553516, 0.006097163539379835, -0.00814123172312975, -0.006955672521144152, 0.014962580054998398, -0.033989936113357544, 0.029668191447854042,
0.0046312748454511166, -0.003355192020535469, -0.007685696706175804, -0.004056015517562628, 0.015219548717141151, 0.011896477080881596, 0.02019539475440979, -0.026584567502141, -0.02576940787434578, -0.00012529043306130916, 0.05312241613864899, 0.024832509458065033, -0.009431915357708931, 0.015091064386069775, -0.053869958966970444, 0.0040326546877622604, -0.010366346687078476, -0.0042399815283715725,
0.02388639748096466, -0.0026587487664073706, -0.0036296811886131763, -0.02735547348856926, 0.0056912703439593315, 0.04361457750201225, 0.001051965169608593, -0.004298383370041847, 0.005577386356890202, 0.0450395829975605, 0.02553333342075348, -0.027635803446173668, -0.04370801895856857, -0.028663678094744682, -0.00201048725284636, -0.0025200441014021635, -0.024949314072728157, -0.020277157425880432,
0.009081503376364708, -0.00627820985391736, -0.01617734134197235, -0.015009301714599133, -0.0050488486886024475, -0.0354616641998291, -0.014904177747666836, -0.04347441345453262, -0.025907104834914207, -0.028476791456341743, -0.012568099424242973, -0.022134339436888695, -0.019809942692518234, -0.053262580186128616, -0.006704544182866812, -0.01973985880613327, -0.009063982404768467, 0.008853735402226448,
0.051393717527389526, 0.016212381422519684, -0.026607928797602654, -0.036349374800920486, -0.006202287040650845, 0.003980093169957399, -0.009274229407310486, 0.009992573410272598, 0.013011954724788666, -0.021048063412308693, -0.006949832197278738, -0.016632875427760372, 0.031116560101509094, -0.02143351547420025, 0.0167730475908335, -0.03291533887386322, -0.01882878877222538, -0.017719151452183723, -
0.059476546943187714, -0.011259895749390125, 0.012603141367435455, -0.010074337013065815, -0.005253255367279053, -0.0027507320046424866, -0.011487663723528385, -0.037353888154029846, 0.03186410292983055, 0.04018054157495499, 0.02597718872129917, -0.05120683088898659, -0.03272845223546028, -0.00289526767606824636, 0.0018761627143248916, -0.011172293685376644, -0.0447826161980629, -0.026724733412265778, -

Succesfull AI implementations

- **Data**
- **Important steps to avoid common mistakes**

Massive fragmentation of the modern tech stack

THE 2024 MAD (MACHINE LEARNING, ARTIFICIAL INTELLIGENCE & DATA) LANDSCAPE



Version 1.0 - March 2024 © Matt Turck (@mattturck) , Aman Kabeer (@AmanKabeer11) & FirstMark (@firstmarkcap) Blog post: mattturck.com/MAD2024 Interactive version: MAD.firstmarkcap.com Comments? Email MAD2024@firstmarkcap.com

FIRSTMARK
EARLY STAGE VENTURE CAPITAL

From Vision to Value: Building Your AI Solution

AI success isn't about building models - it's about aligning data, governance, and business value.

1 **Start with your data**
Transformation, Integration, Storage, Governance, security, classification, access

2 **Define the Business Need**
Identify repetitive, high-impact processes (HR, IT, support). Align goals with measurable outcomes.

3 **Evaluate Existing Tools**
Explore Azure AI Foundry, and pre-trained models before custom development.

4 **Choose the Right Architecture**
Combine Azure AI Search, Blob Storage, and RAG pipelines to keep data secure within your tenant.

5 **Pilot, Measure, Scale**
Launch a controlled pilot, gather feedback, optimize, then expand to broader enterprise use cases.

We are a leading Microsoft partner

We specialize in all Microsoft Cloud Solution Areas. Our customers can rely on our comprehensive portfolio of Microsoft Cloud services, supported by all 6 Solution Designations, alongside our own services for digital transformation, management, security and modernization

25+
Years of collaboration

600+
Microsoft Certified Professionals

One of only several
Microsoft globally-managed partners worldwide

Best licensing practices
LSP authorization in 36 countries
CSP authorization in 50+ countries
SPLA reseller for 8 years

Microsoft Partner of the Year


Oman


Bulgaria


Cambodia x2


Malaysia


Vietnam x2

2021
2022
2024

Microsoft Solutions Partner
Microsoft Cloud

Microsoft Solutions Partner
Infrastructure
Azure

Microsoft Solutions Partner
Data & AI
Azure

Microsoft Solutions Partner
Digital & App Innovation
Azure

Microsoft Solutions Partner
Security

Microsoft Solutions Partner
Modern Work

Microsoft Solutions Partner
Business Applications

15 Advanced Specializations

Microsoft Partner | Azure Expert MSP

Since 2019

Azure

- Infra and Database Migration to Microsoft Azure
- Microsoft Azure Virtual Desktop
- Kubernetes on Microsoft Azure
- DW Migrate
- Azure Analytics

Security

- Cloud Security
- Identity and Access Management
- Threat Protection
- Information Protection and Governance

Modern Work & Business Apps

- Adoption and Change Management
- Calling for Microsoft Teams
- Teamwork Deployment
- Meetings & Meeting Rooms for Teams
- Modernize Endpoints
- Microsoft Low Code Application Development

NOVENTIQ
Global expertise, local outcomes