

MODERN PROBLEMS REQUIRE MODERN SOLUTIONS

**FINDING YOUR MEME TWIN
WITH EMBEDDINGS & VECTOR DATABASES**

Redis



Guy Royse
Developer Advocate

Redis



@guy.dev



@guyroyse



github.com/guyroyse



guy.dev

I DON'T ALWAYS PRETEND TO BE A DATA SCIENTIST



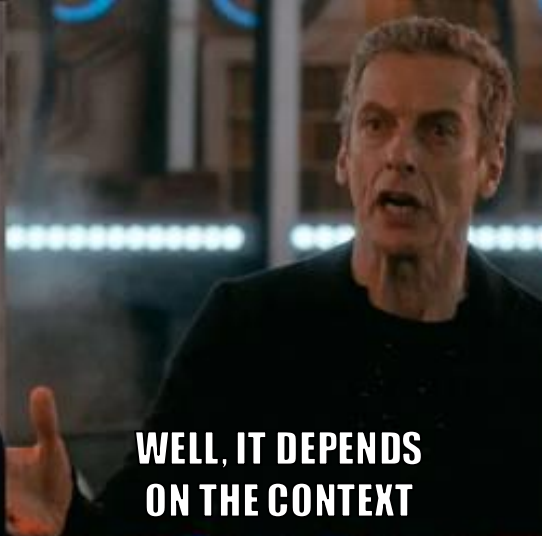
BUT WHEN I DO, I DO IT ON STAGE



I'M SOMETHING OF A DATA SCIENTIST MYSELF



**IS FOUR AGENDA
ITEMS A LOT?**



**WELL, IT DEPENDS
ON THE CONTEXT**



**FOR A TALK AT
A CONFERENCE? NO**



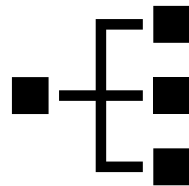
**FOR A TALK WITH
YOUR BOSS? YES**

```
[ 0.86 -0.72 0.25  
-0.93 -0.13 -0.64  
0.47 0.39 0.78  
-0.58 -0.41 0.63 ]
```

What's an embedding?



Using
embeddings



Making
embeddings



Code & demo

A man with a beard and sunglasses, wearing a dark shirt and a light-colored vest, is holding a handgun with both hands, aiming it upwards. He is in a shooting range, with a blurred background showing other people and equipment. The text "AM I THE ONLY ONE AROUND HERE WHO KNOWS WHAT AN EMBEDDING IS?" is overlaid on the bottom of the image in a bold, white, sans-serif font.

**AM I THE ONLY ONE AROUND HERE WHO KNOWS
WHAT AN EMBEDDING IS?**

Structured data



Numbers



Strings



Locations

Unstructured data



Text



Images



Audio



**STORING ALL YOUR COMPANY'S
STRUCTURED DATA IN **SQLITE****

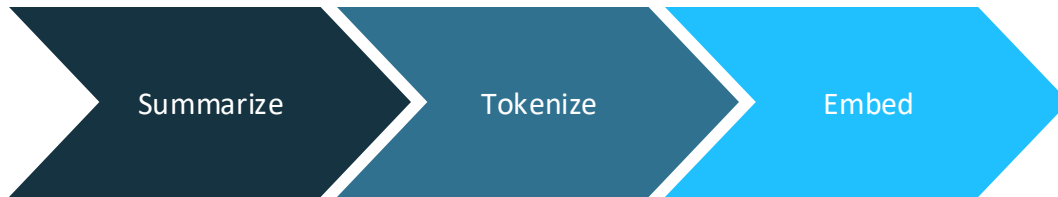


**STORING ALL YOUR COMPANY'S
STRUCTURED DATA IN **ACCESS****



**STORING ALL YOUR COMPANY'S
STRUCTURED DATA IN **EXCEL****

DON'T NEED TO
DEBUG YOUR
CODE IF IT
DOESN'T
COMPILE



[0.86 -0.72 0.25
-0.93 -0.13 -0.64
0.47 0.39 0.78
-0.58 -0.41 0.63]



[0.72 0.86 -0.39
-0.63 0.95 -0.21
-0.86 0.61 0.45
-0.79 -0.03 0.70]



**HAVING
UNSTRUCTURED
DATA**



**HAVING
STRUCTURED
DATA**



**USING AI TO GET
FEATURES
FROM
UNSTRUCTURED
DATA**

+1.0

0

-1.0

Elderly

Alcoholic





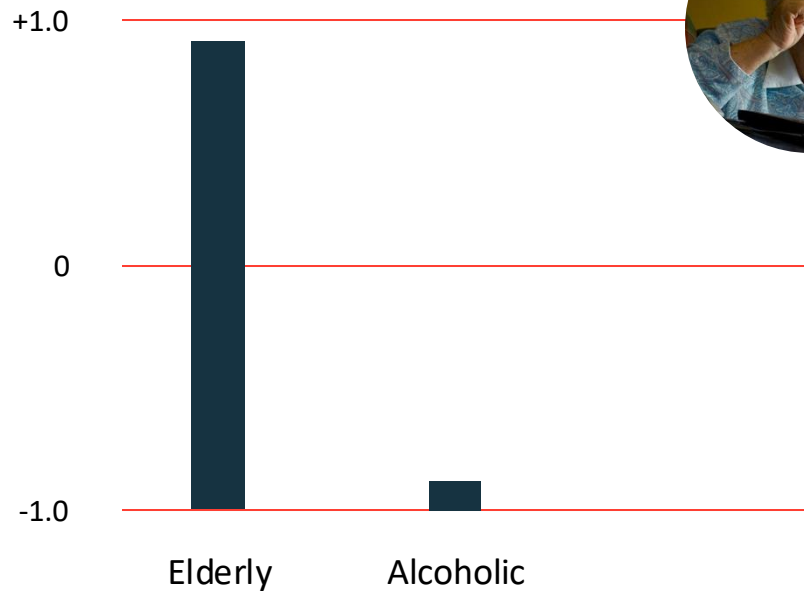
**HAVING
UNSTRUCTURED
DATA**

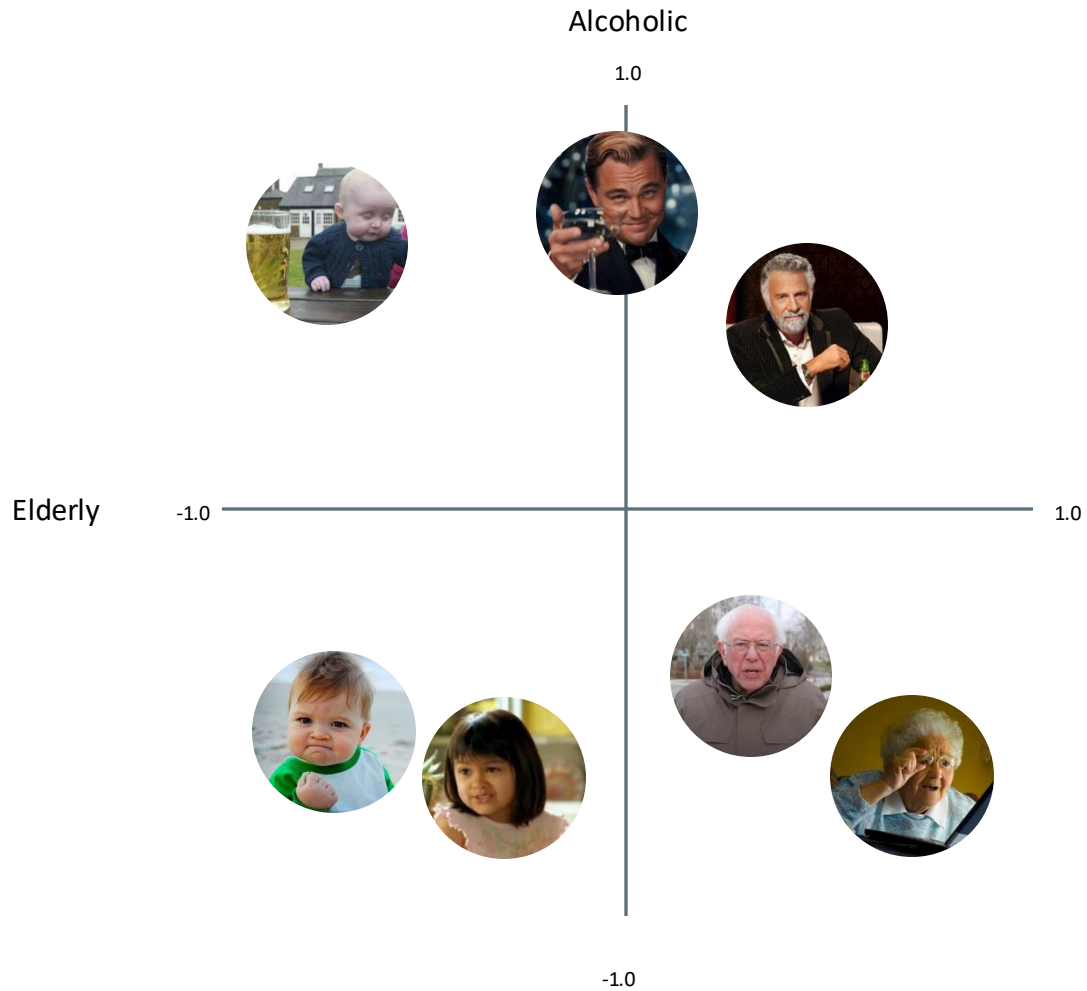


**HAVING
STRUCTURED
DATA**



**USING AI TO GET
FEATURES
FROM
UNSTRUCTURED
DATA**

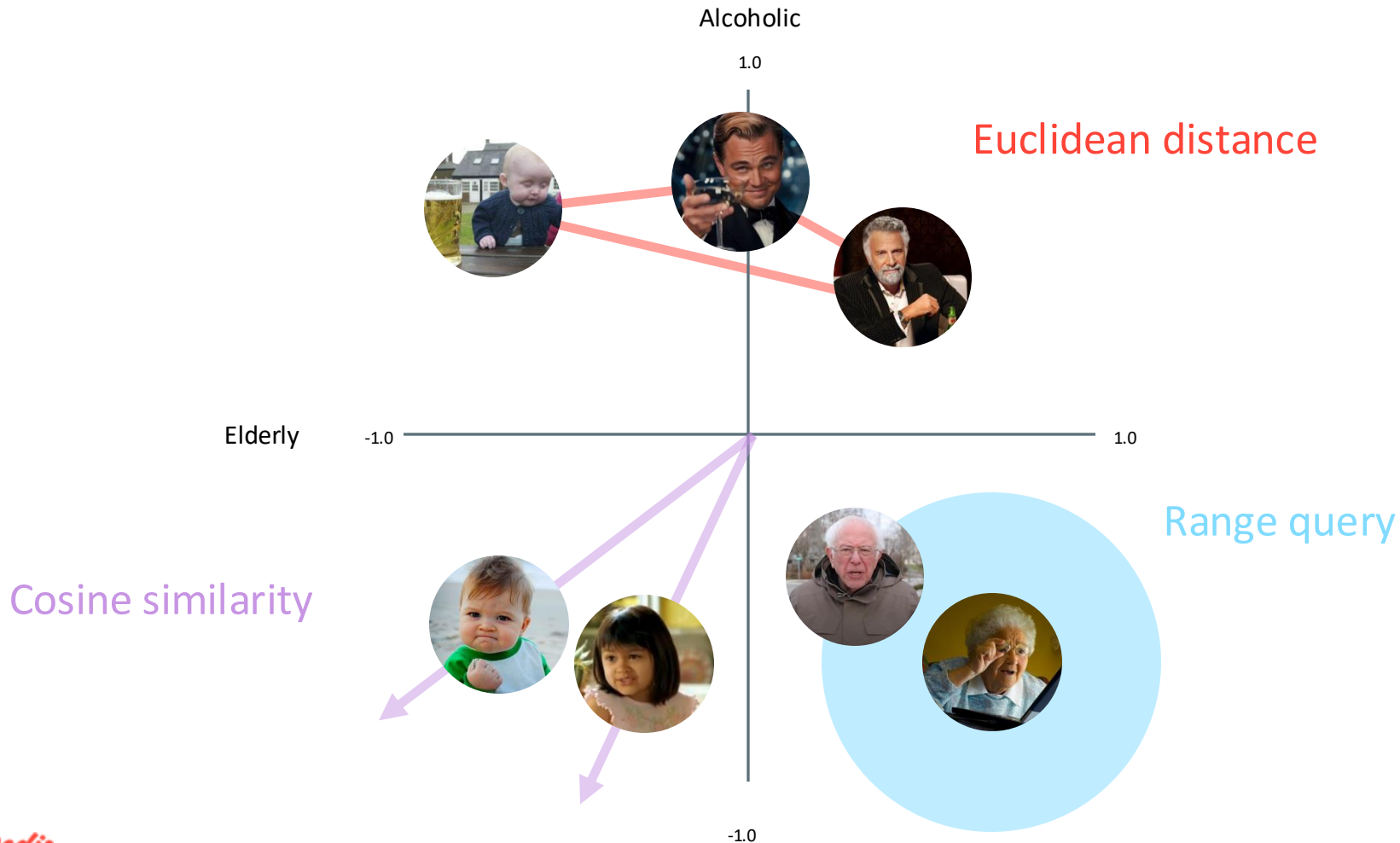





$$\begin{bmatrix} 0.72 & 0.86 & -0.39 \\ -0.63 & 0.95 & -0.21 \\ -0.86 & 0.61 & 0.45 \\ -0.79 & -0.03 & 0.70 \end{bmatrix}$$

$$\begin{bmatrix} 0.86 & -0.72 & 0.25 \\ -0.93 & -0.13 & -0.64 \\ 0.47 & 0.39 & 0.78 \\ -0.58 & -0.41 & 0.63 \end{bmatrix}$$
$$\begin{bmatrix} 0.89 & -0.76 & 0.23 \\ -0.90 & -0.03 & -0.68 \\ 0.45 & 0.39 & 0.68 \\ -0.42 & -0.40 & 0.57 \end{bmatrix}$$

$$\begin{bmatrix} -0.96 & 0.52 & -0.77 \\ 0.68 & -0.95 & 0.34 \\ 0.82 & -0.29 & 0.17 \\ -0.43 & -0.61 & 0.79 \end{bmatrix}$$





Flat

- A simple brute-force search
- Perfect accuracy
- Runs into scaling limitations

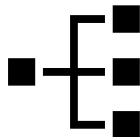


HNSW

- Hierarchical navigable small world
- Creates layers of graphs in a hierarchy
- Approximate accuracy
- Tunable to balance accuracy, speed, and space

SEMANTIC SEARCH

DON'T NEED TO
DEBUG YOUR CODE IF
IT DOESN'T COMPILE



[0.89 -0.76 0.23
-0.90 -0.03 -0.68
0.45 0.39 0.68
-0.42 -0.40 0.57]

DELETION IS 100%
FILE COMPRESSION

TYPESCRIPT IS JUST
JAVASCRIPT FOR C#
DEVS CHANGE MY
MIND

CODE DOESN'T
COMPILE. COMPILES
IT AGAIN.



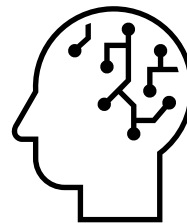
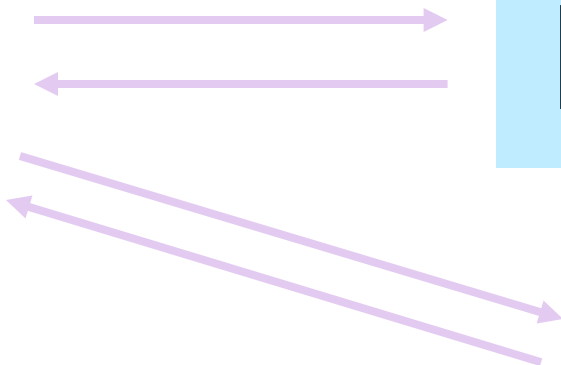
SEMANTIC CACHING



Can you explain to me
what a dank meme is?

[0.89 -0.76 0.23
-0.90 -0.03 -0.68
0.45 0.39 0.68
-0.42 -0.40 0.57]

]

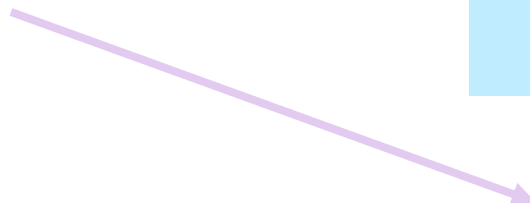


RETRIEVAL AUGMENTED GENERATION



Can you explain to me
what a dank meme is?

[0.89 -0.76 0.23
-0.90 -0.03 -0.68
0.45 0.39 0.68
-0.42 -0.40 0.57]



Overused
memes are dank

Dank memes are
the dankest

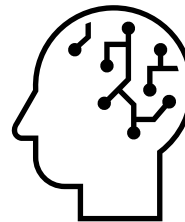
A dank meme is
exceptionally odd or
unique

Given the following context:



Answer the following question:

Can you explain to me
what a dank meme is?

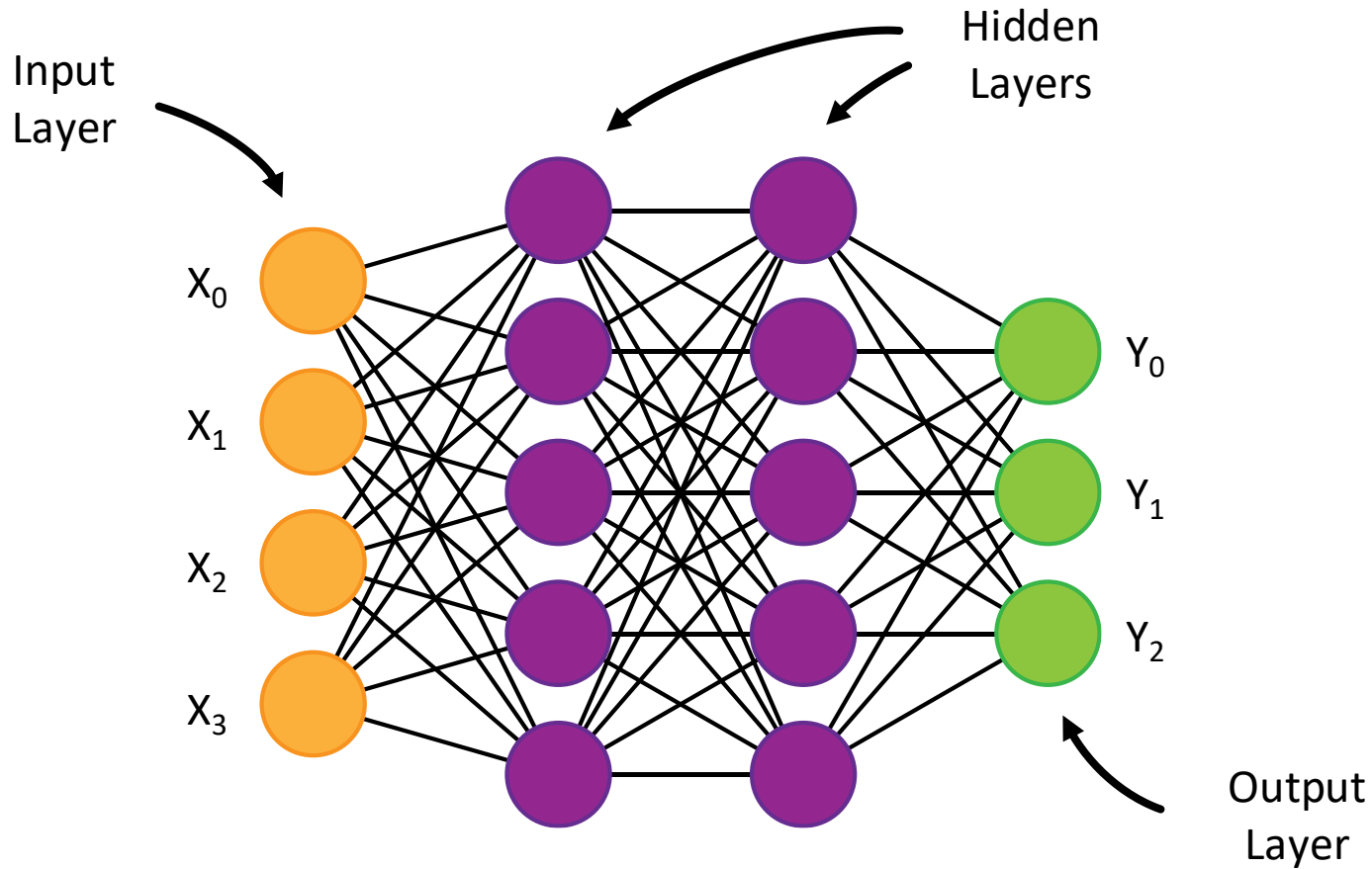




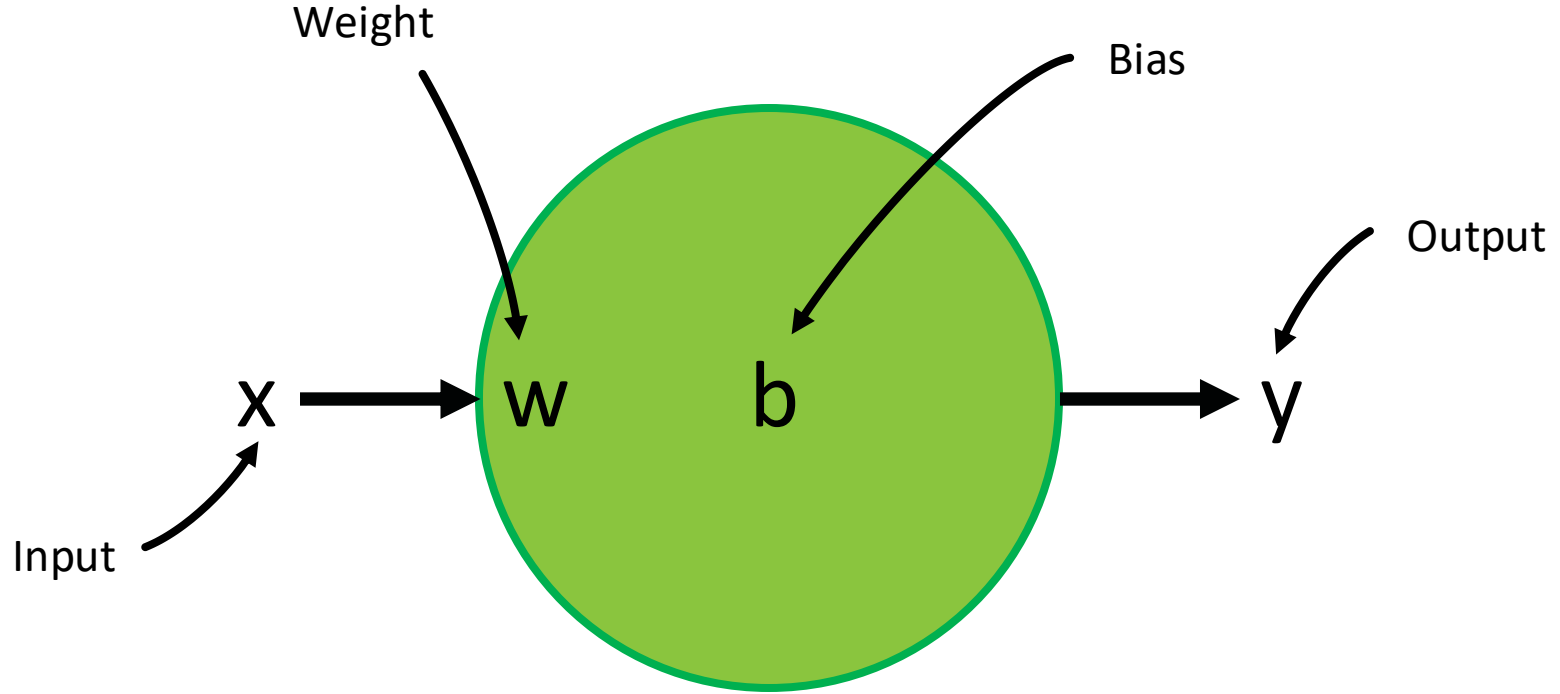
**CREATING EMBEDDING
MODELS YOURSELF**



**USING EXISTING
EMBEDDING MODELS**



YOUR BASIC NEURAL NETWORK



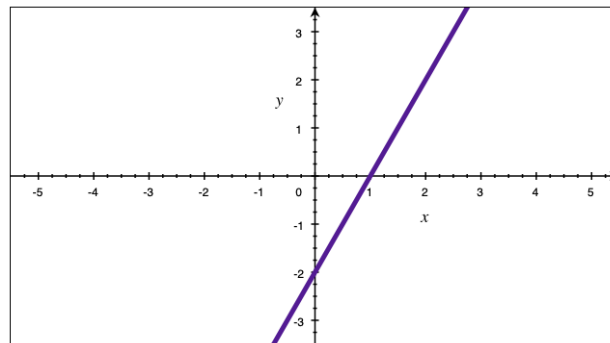
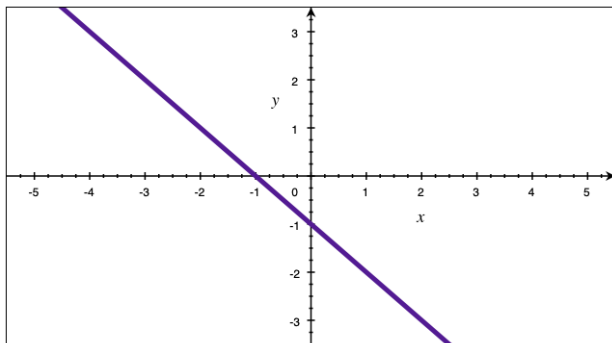
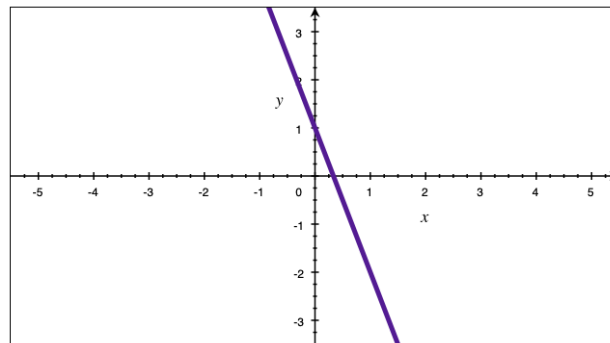
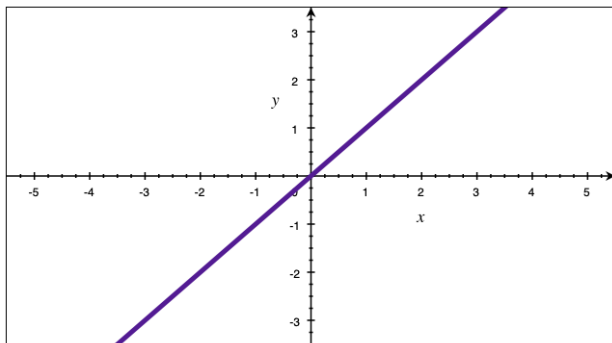
YOUR BASIC NEURON

Diagram illustrating the linear equation $xw + b = y$. The components are labeled as follows:

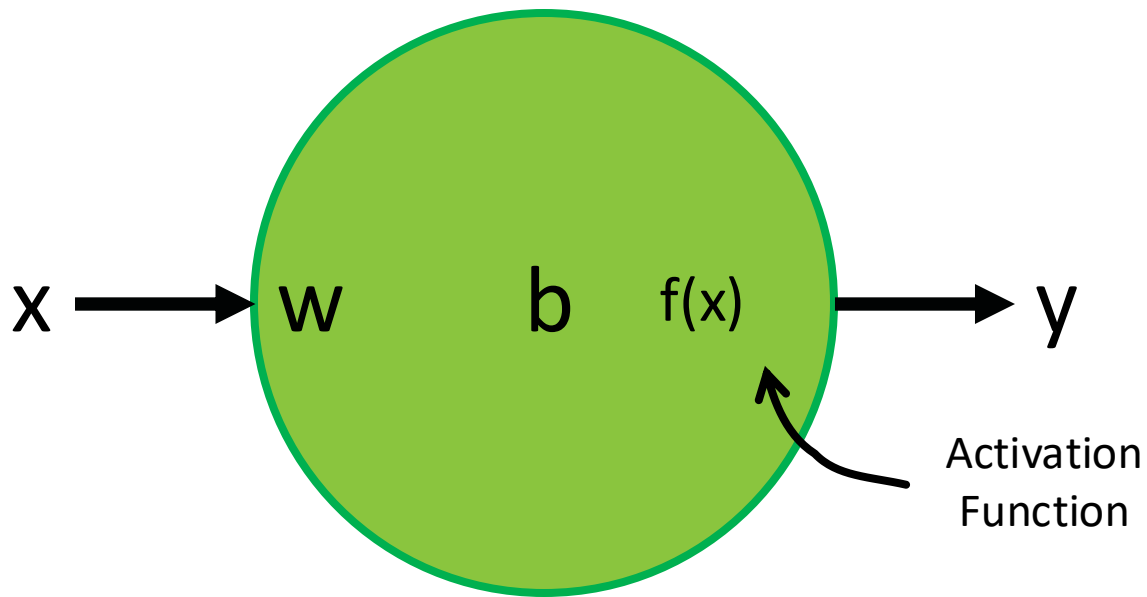
- x : Input
- w : Weight
- b : Bias
- y : Output

Diagram illustrating the linear equation $y = mx + b$. The components are labeled as follows:

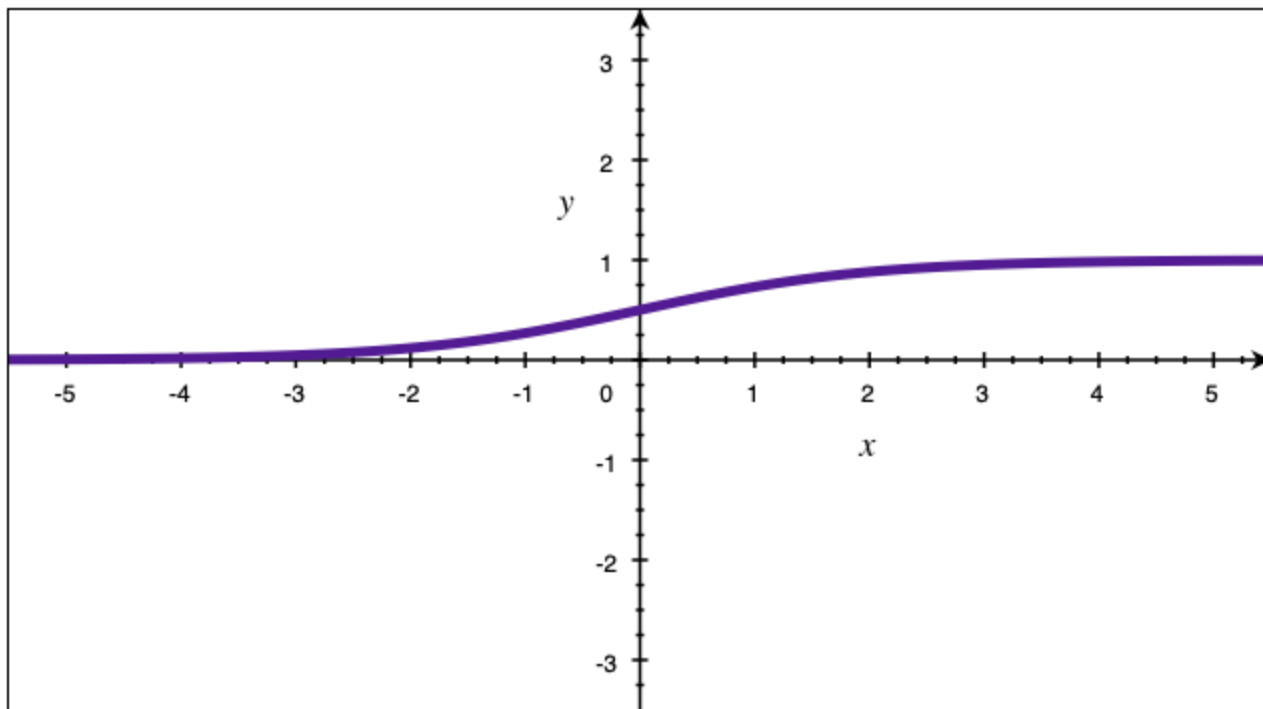
- m : Slope
- b : Y-Intercept



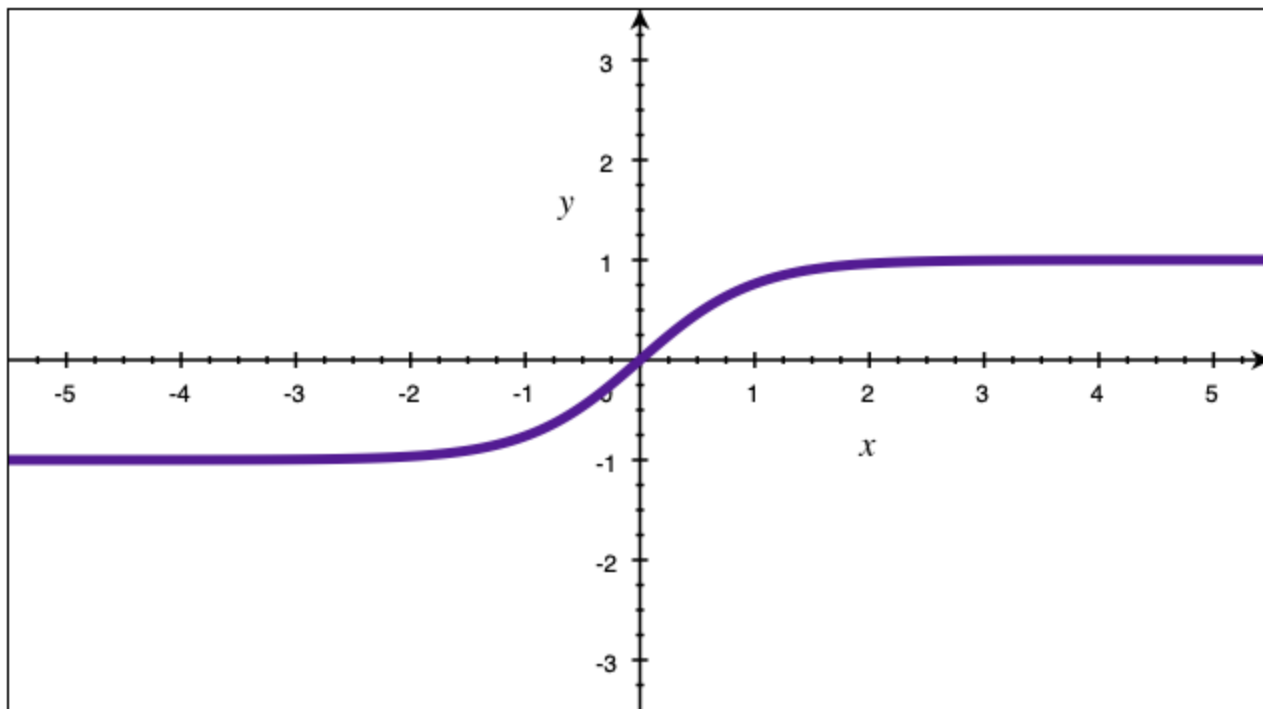
LINEAR FUNCTIONS



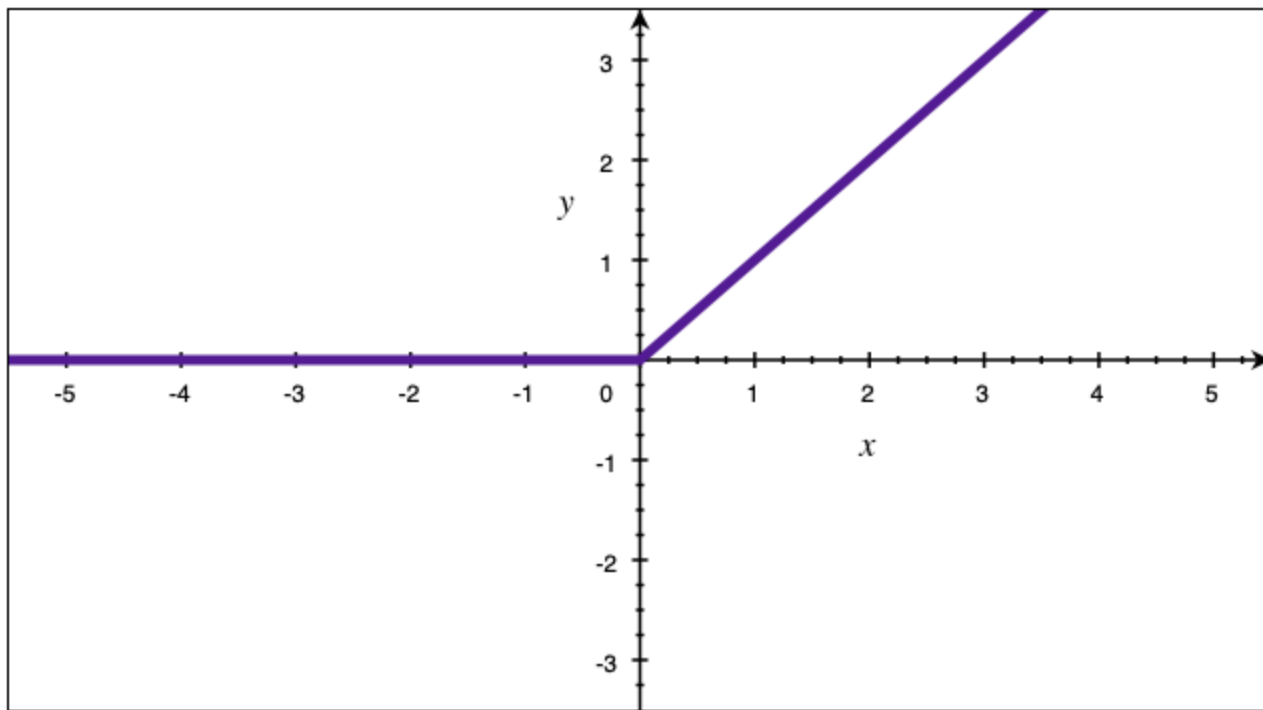
LESS LINEAR FUNCTIONS



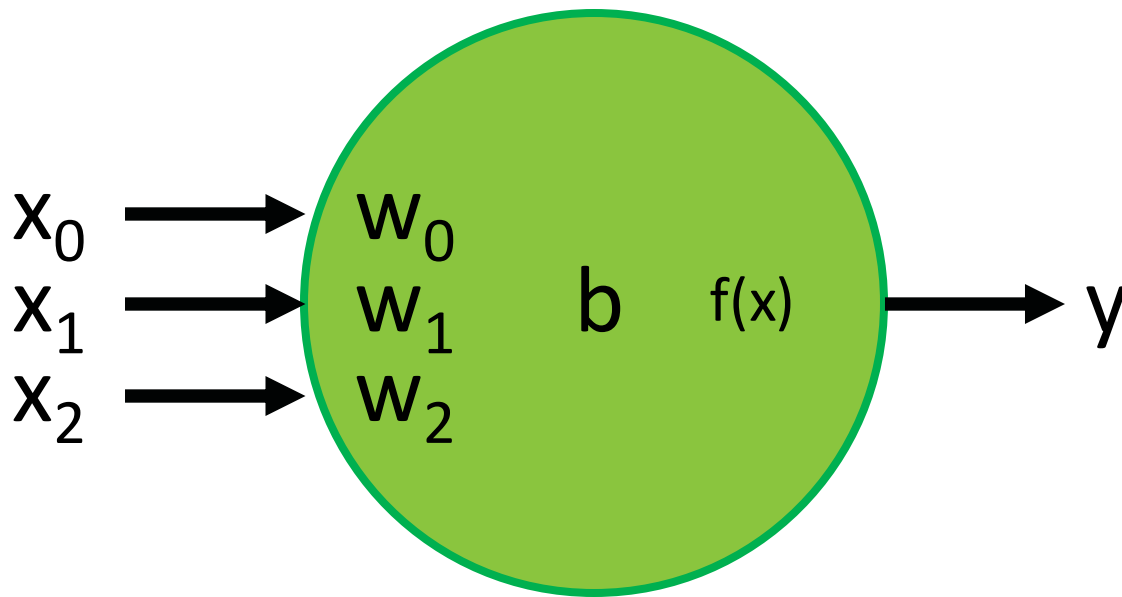
SIGMOID



HYPERBOLIC TANGENT



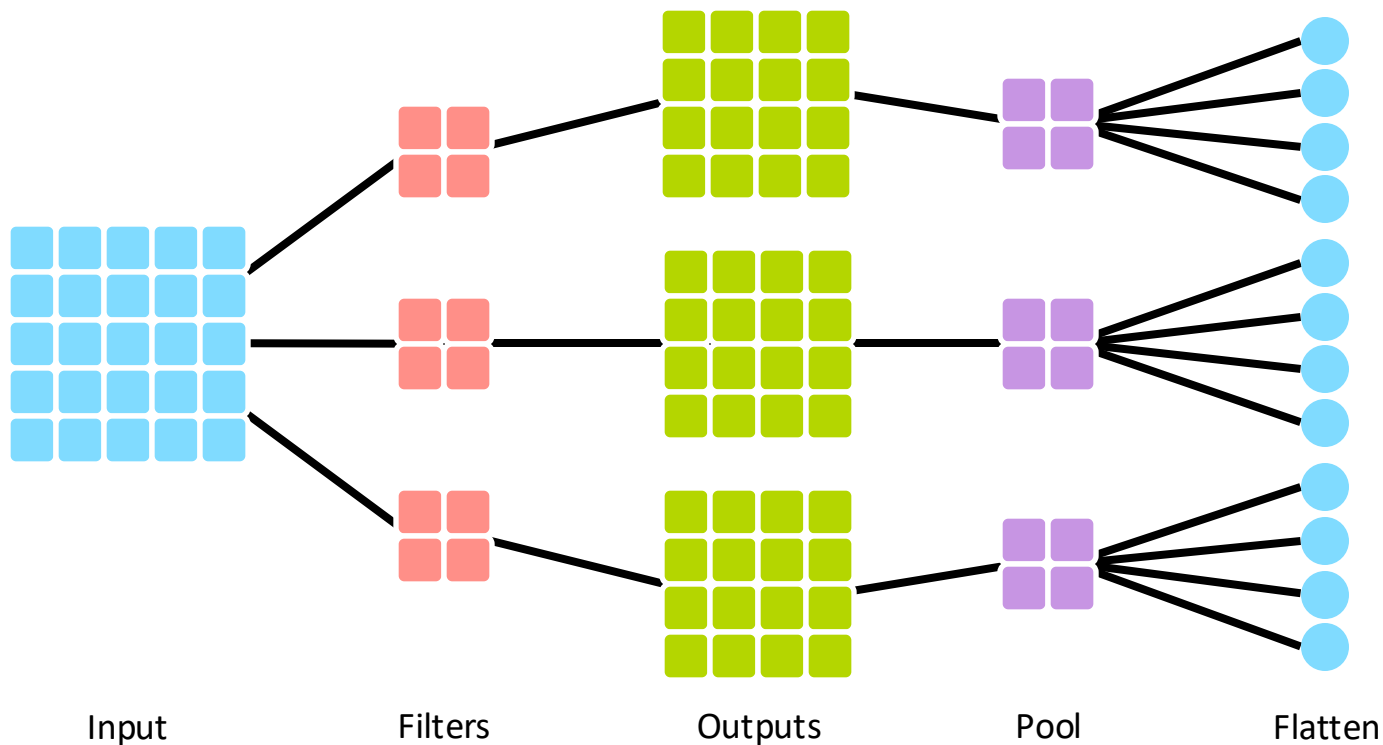
RECTIFIED LINEAR UNITS (ReLU)



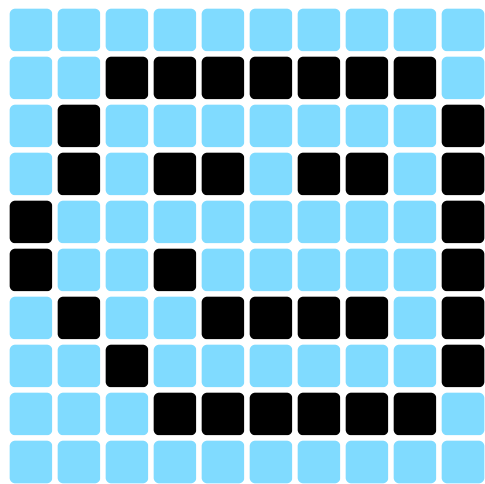
A MORE REALISTIC NEURON

$$x_0w_0 + x_1w_1 + x_2w_2 + b = y$$

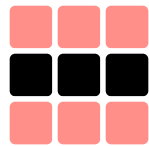
$$\sum_i x_iw_i + b$$



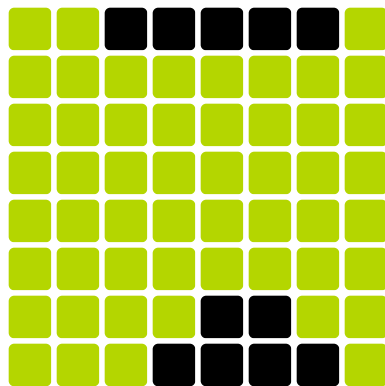
CONVOLUTION NEURAL NETWORK



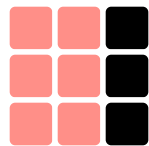
Input



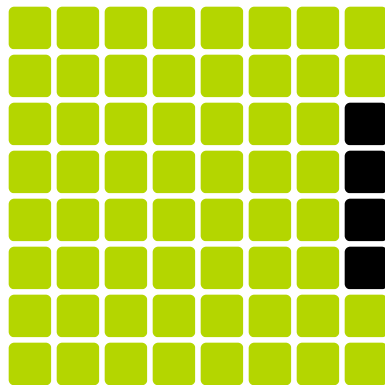
Filter



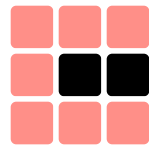
Output



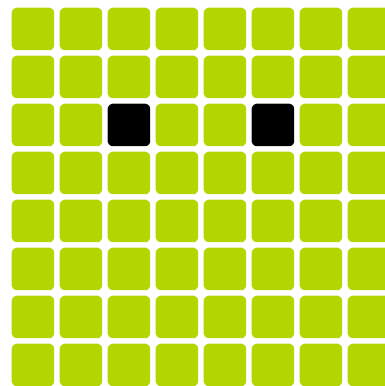
Filter



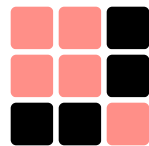
Output



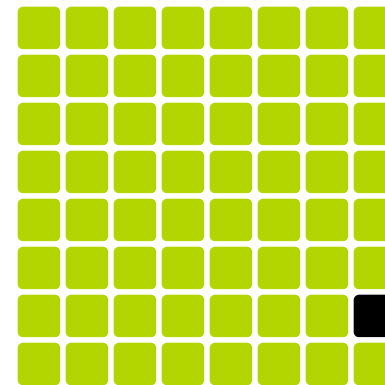
Filter



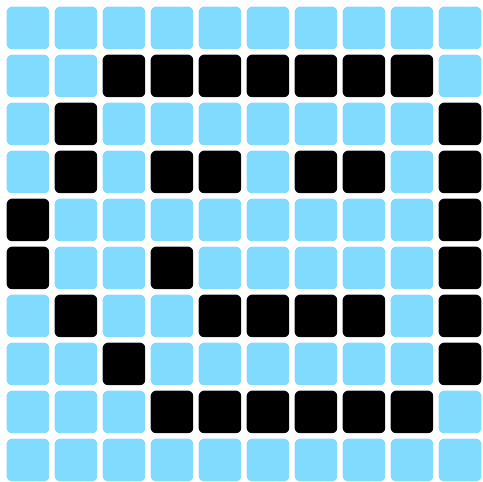
Output



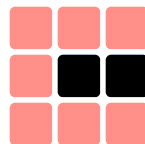
Filter



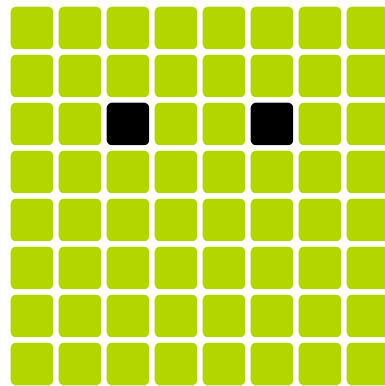
Output



Input

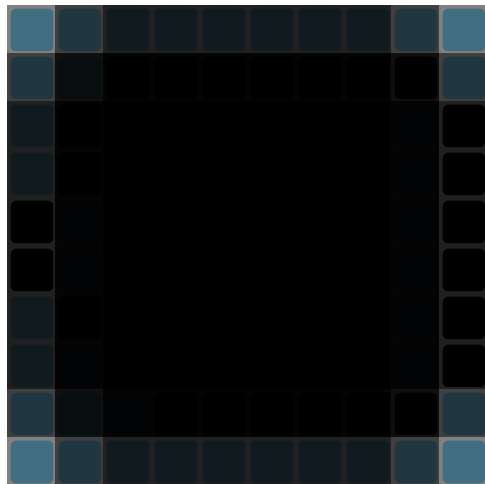


Filter

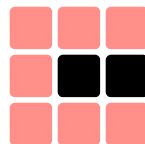


Output

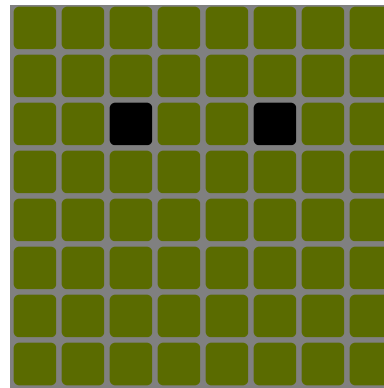
CONVOLUTIONS



Input

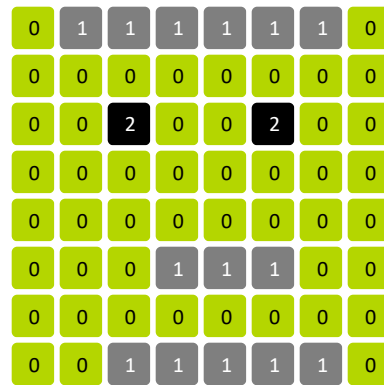
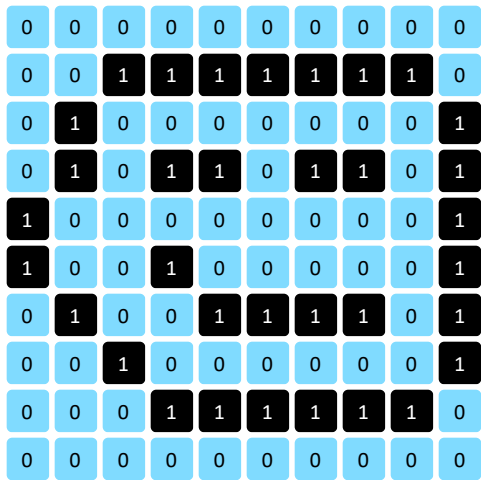


Filter



Output

CONVOLUTIONS



CONVOLUTIONS WITH MATH

x_i

0	0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	1	1	0
0	1	0	0	0	0	0	0	0	1
0	1	0	1	1	0	1	1	0	1
1	0	0	0	0	0	0	0	0	1
1	0	0	1	0	0	0	0	0	1
0	1	0	0	1	1	1	1	0	1
0	0	1	0	0	0	0	0	0	1
0	0	0	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0	0	0

 w_i

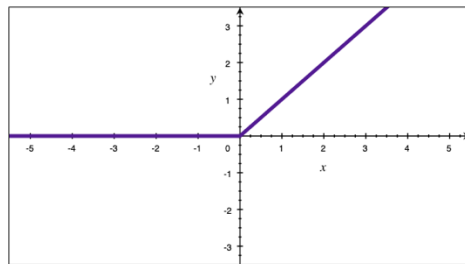
-1	-1	-1
-1	1	1
-1	-1	-1
0		

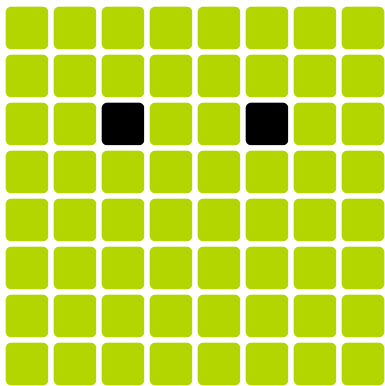
 b

$$\sum_i x_i w_i + b$$

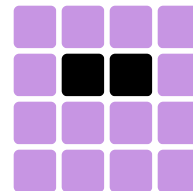
 y_i

0	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0
0	0	2	0	0	2	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0
0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	0

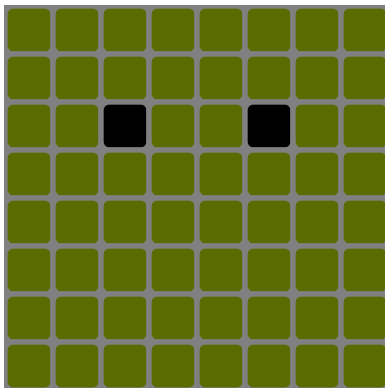




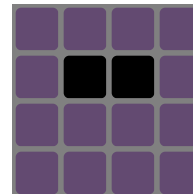
2x2



POOLING



2x2



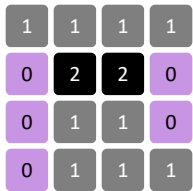
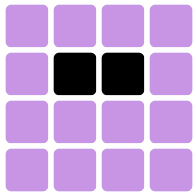
POOLING

0	1	1	1	1	1	1	0
0	0	0	0	0	0	0	0
0	0	2	0	0	2	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	1	1	1	0	0
0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	0

2x2

1	1	1	1
0	2	2	0
0	1	1	0
0	1	1	1

POOLING



FLATTENING

**CUSTOM
EMBEDDING
MODELS**

DEVS

**EXISTING
EMBEDDING
MODELS**



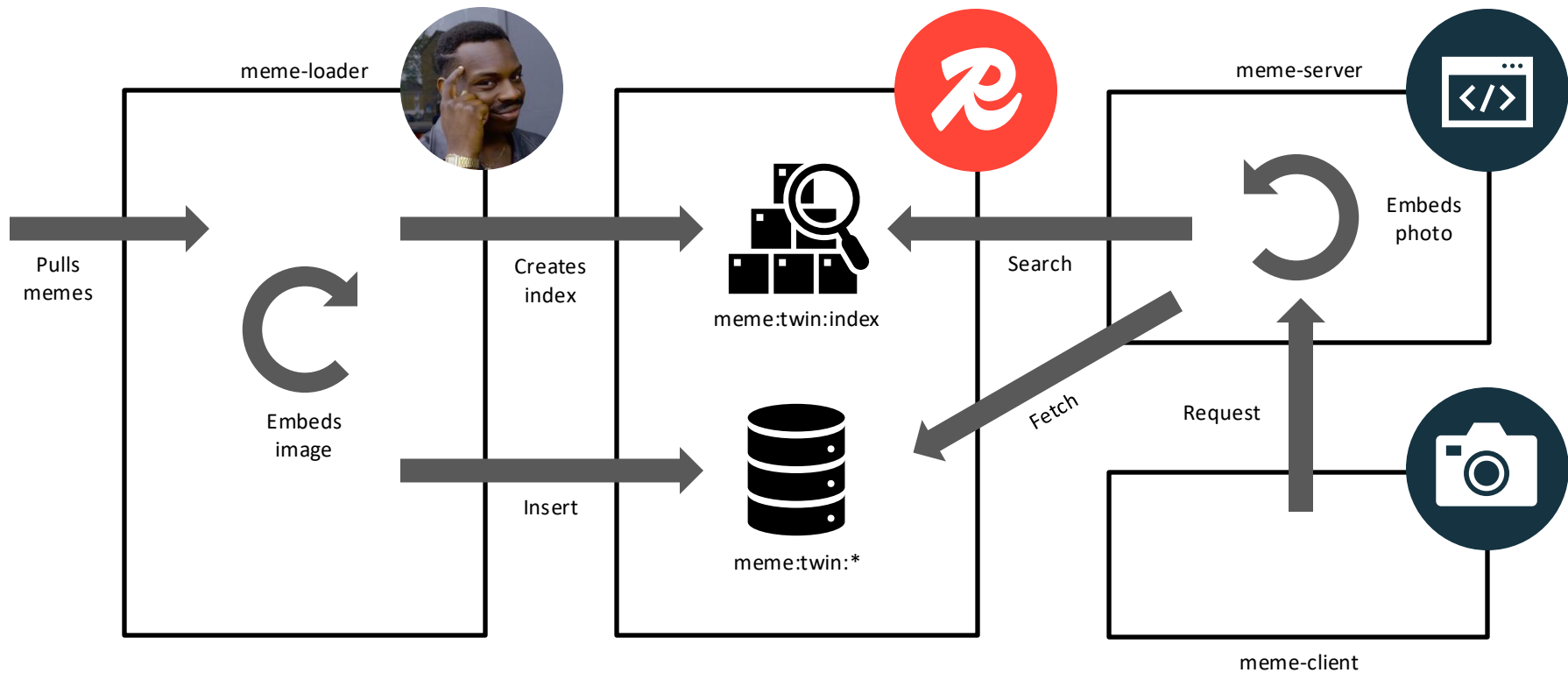
**ONE DOES NOT SIMPLY
CREATE THEIR OWN EMBEDDING MODELS**



**Y'ALL GOT ANY MORE OF THOSE
VECTOR DATABASE DEMOS**

I'M NOT SAYING I'M GONNA USE REDIS

BUT I'M GONNA USE REDIS





ME WHEN I SEE A TALK BUT BEFORE THE DEMO



[github.com/guyroyse/
modern-problems-require-modern-solutions](https://github.com/guyroyse/modern-problems-require-modern-solutions)

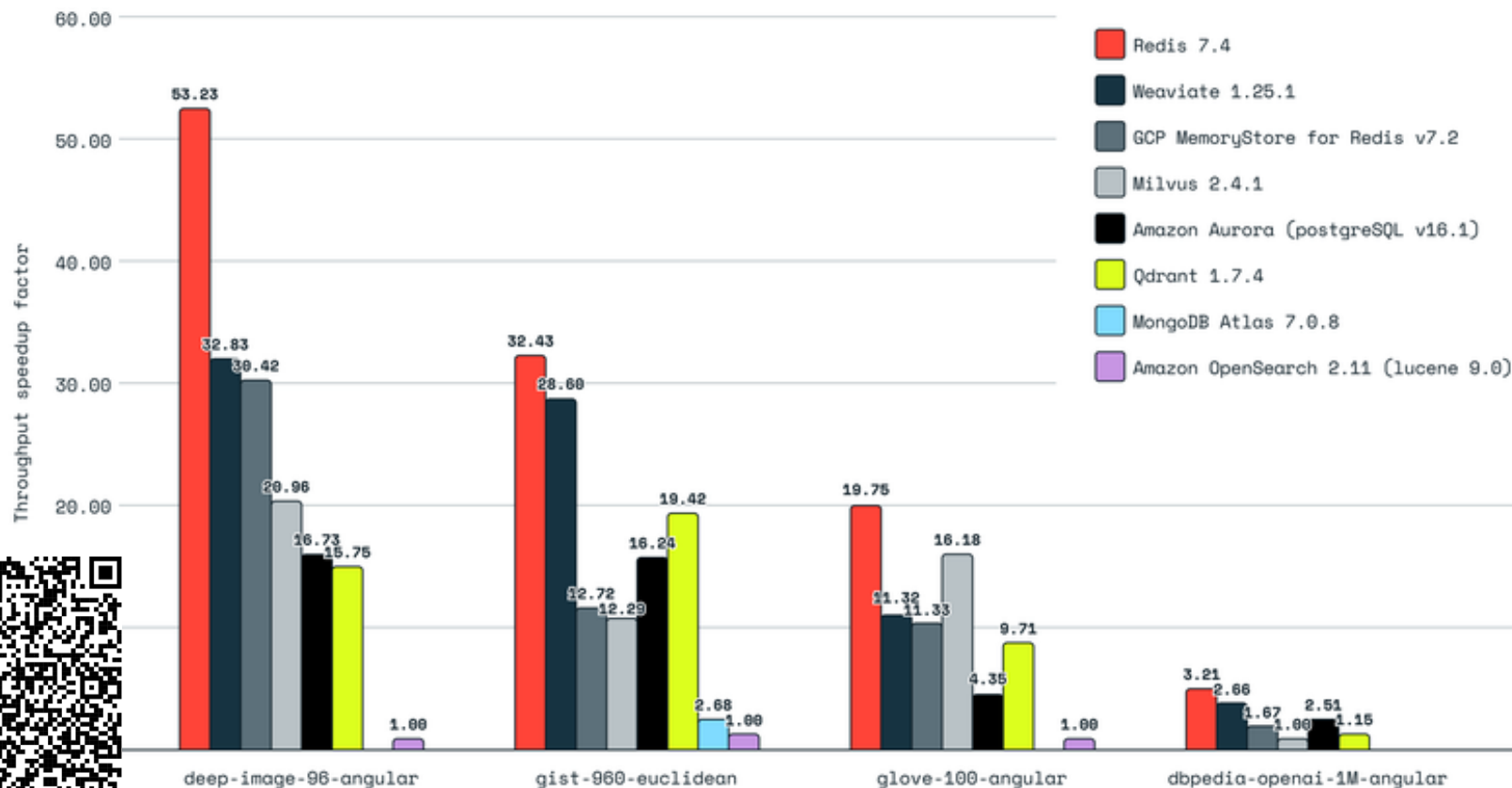


redis.io
university.redis.io



discord.gg/redis

Redis offers better performance than any other vector database





Guy Royse
Developer Advocate

Redis



@guy.dev



@guyroyse



github.com/guyroyse



guy.dev

MODERN PROBLEMS REQUIRE MODERN SOLUTIONS

**FINDING YOUR MEME TWIN
WITH EMBEDDINGS & VECTOR DATABASES**

Redis