

Beginner's Guide to Vector Databases

AI by Hand 🖋️

Prof. Tom Yeh

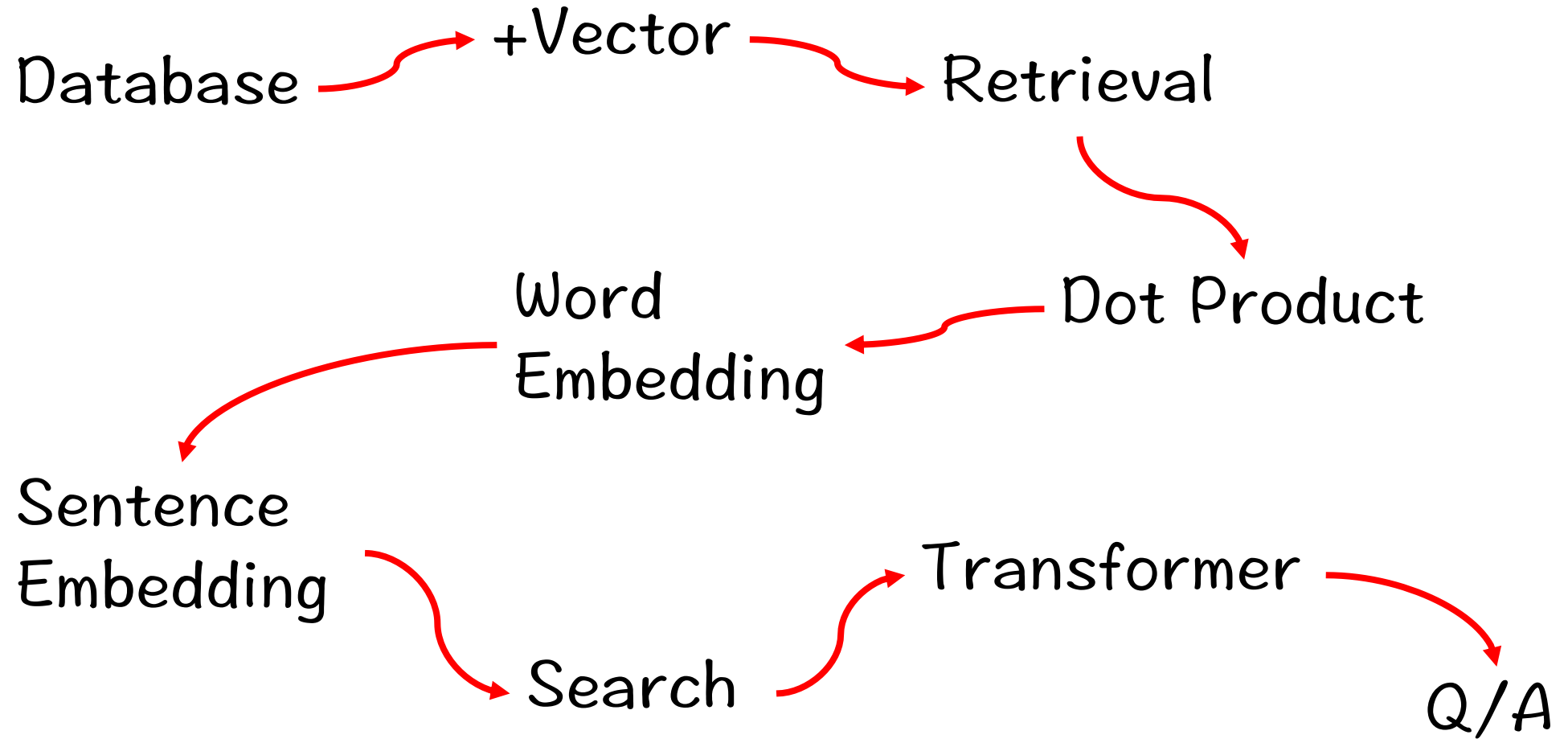


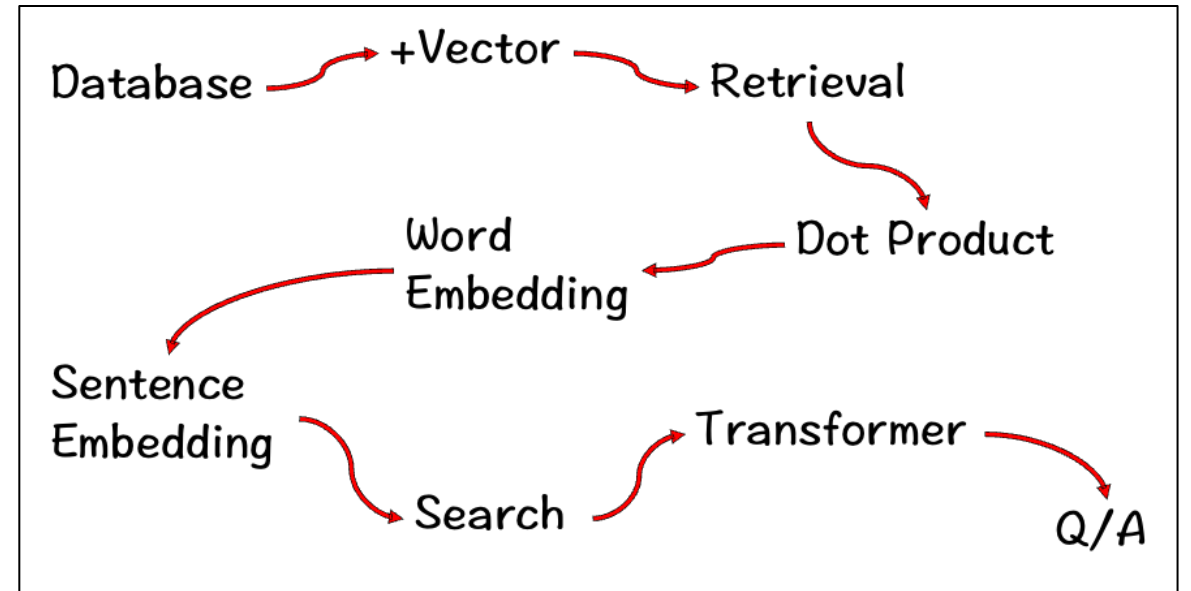
University of Colorado
Boulder

Hosted by:



Roadmap





Database

Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

Fun fact

There are 900 millions dogs in the world!

How to create a table?

SQL: CREATE TABLE animals
(id INT,
name VARCHAR(10),
size INT,
pop INT)

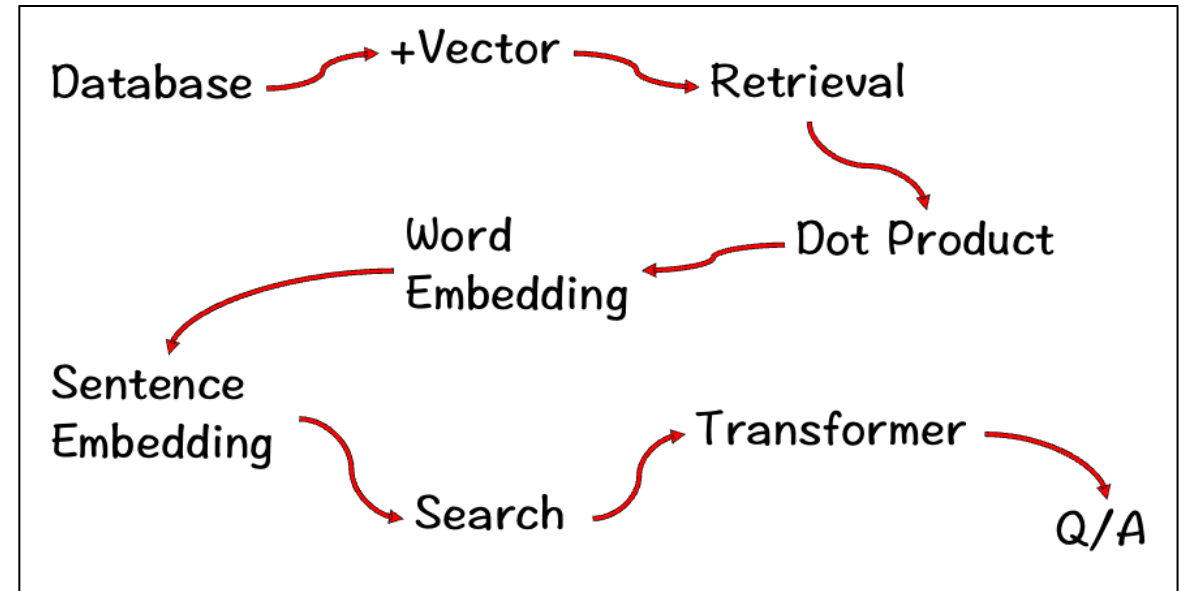
id	name	size	pop

How to insert a record?

SQL:

INSERT INTO animals
VALUES (1, dog, 2, 900)

id	name	size	pop
1	dog	2	900



Vector Database

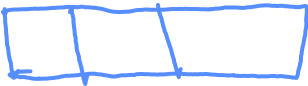
Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

How to create a vector database?

SQL: `CREATE TABLE animals`
`(id INT,`
`name VARCHAR(10),`
`size INT,`
`pop INT,`
`emb VECTOR(3) not null)`

id	name	size	pop	emb
				

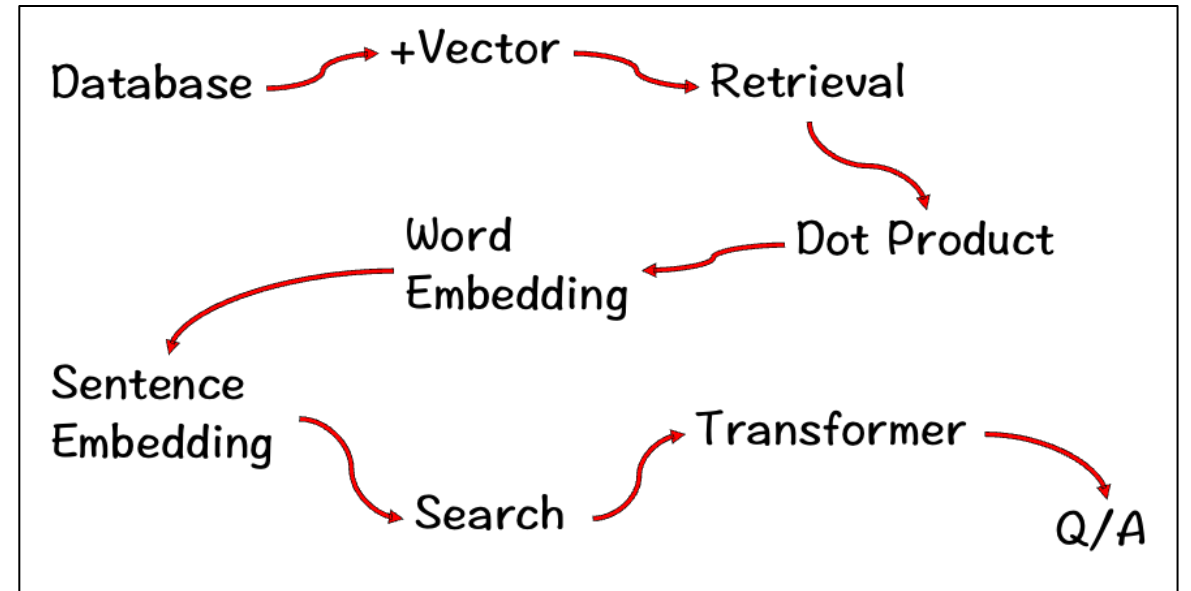
How to insert a record with a vector?

SQL:

INSERT INTO animals

VALUES (1, dog, 2, 900, [2, 1, 0])

id	name	size	pop	emb			
1	dog	2	900	<table><tr><td>2</td><td>1</td><td>0</td></tr></table>	2	1	0
2	1	0					



Retrieval

Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

Which record is relevant to the query “cat”?

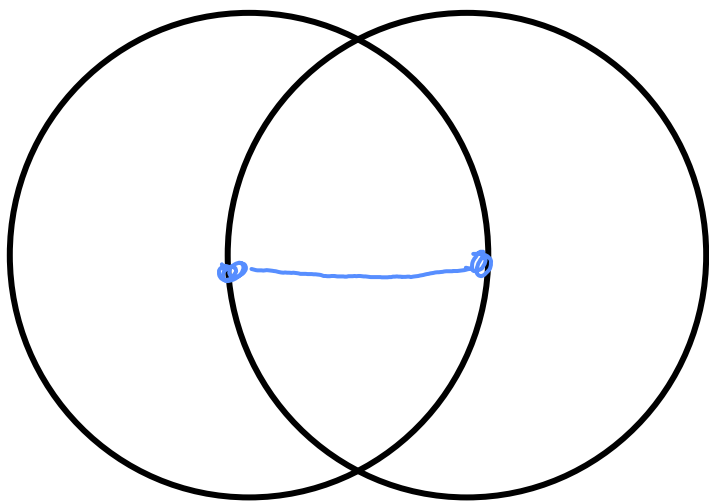
Query

cat			
1	2	0	

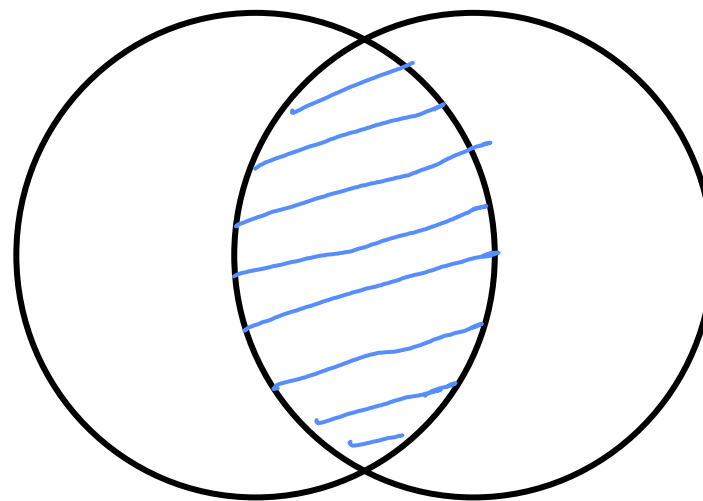
id	name	size	pop	emb			
1	dog	2	900	<table><tr><td>2</td><td>1</td><td>0</td></tr></table>	2	1	0
2	1	0					
2	bat	1	10000	<table><tr><td>0</td><td>1</td><td>2</td></tr></table>	0	1	2
0	1	2					

Draw distance vs similarity

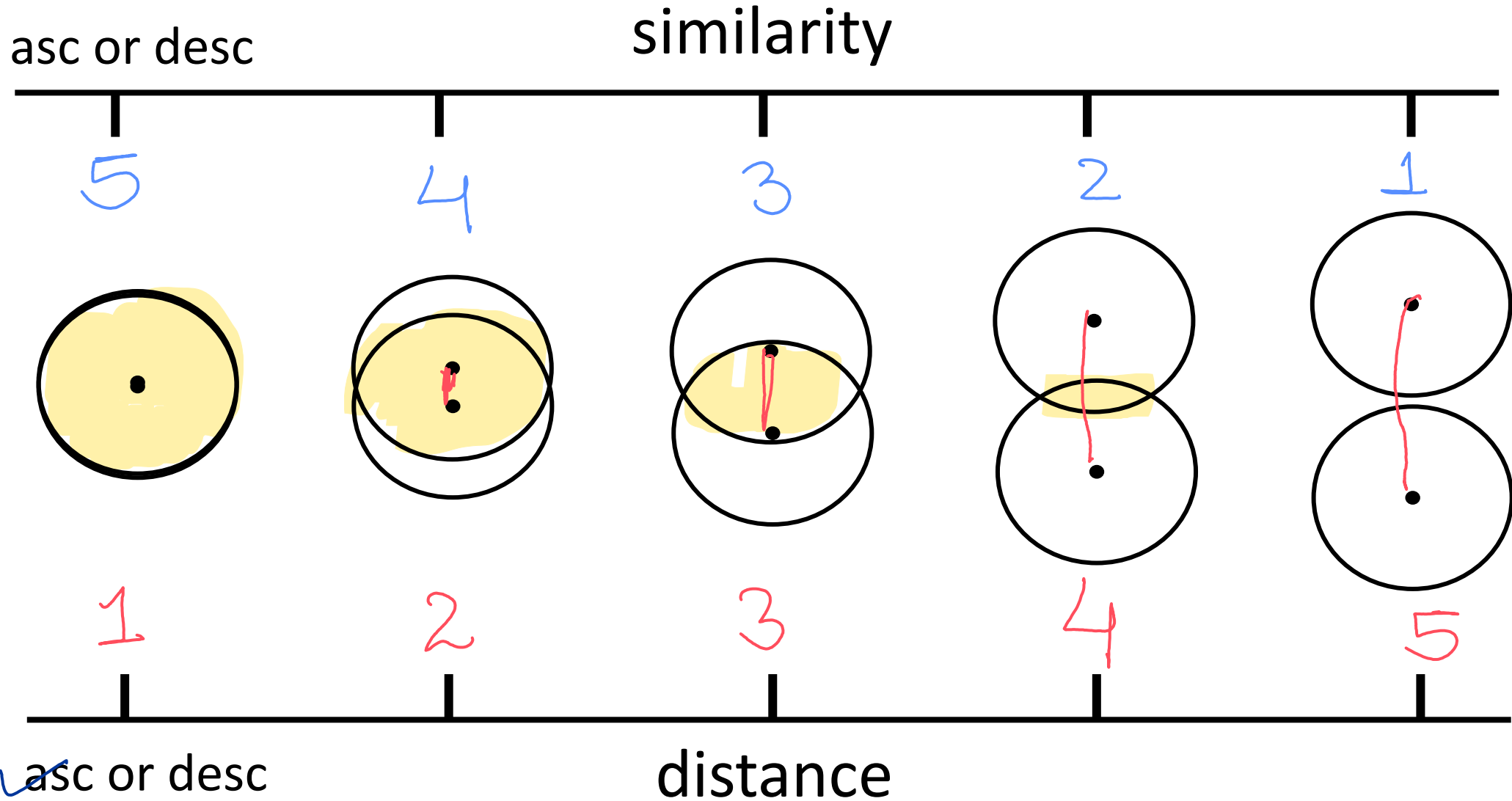
distance



similarity



Distance vs similarity on a scale of 1 to 5



How to retrieve by similarity? (dot product)
↳ descending

SELECT name, emb<___>[1,2,0] AS score

FROM animals

ORDER BY score ~~ASC~~ | DESC ;|

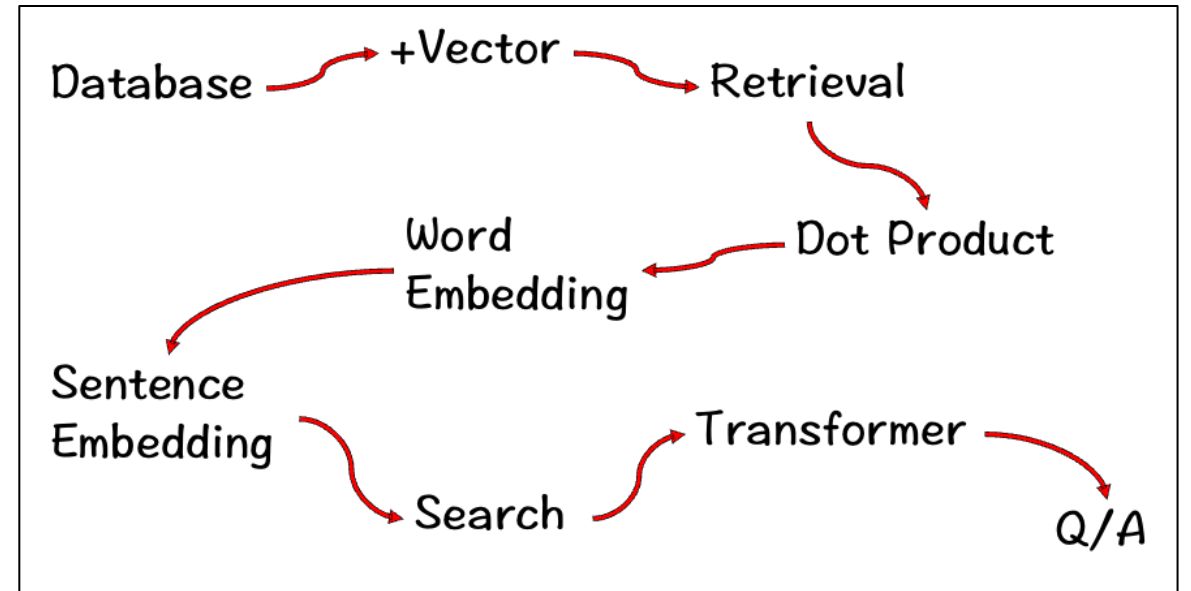
How to retrieve by distance? (Euclidean) L2

~~-~~ → minus

SELECT name, emb<~~*~~>[1, 2, 0] AS score

FROM animals

ORDER BY score ~~DESC~~; ASC



Dot Product

Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

How to compute dot product?

Example:

A

1	2	3
---	---	---

* * *

B

2	2	0
---	---	---

= = = Σ

2	4	0
---	---	---

6

Result

dog

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

* * *

cat

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

= = = Σ

2	2	0
---	---	---

4

Result

similarity

How to compute dot product using matrix multiplication? *scalable*

Example:

			1
			2
			3
2	2	0	6

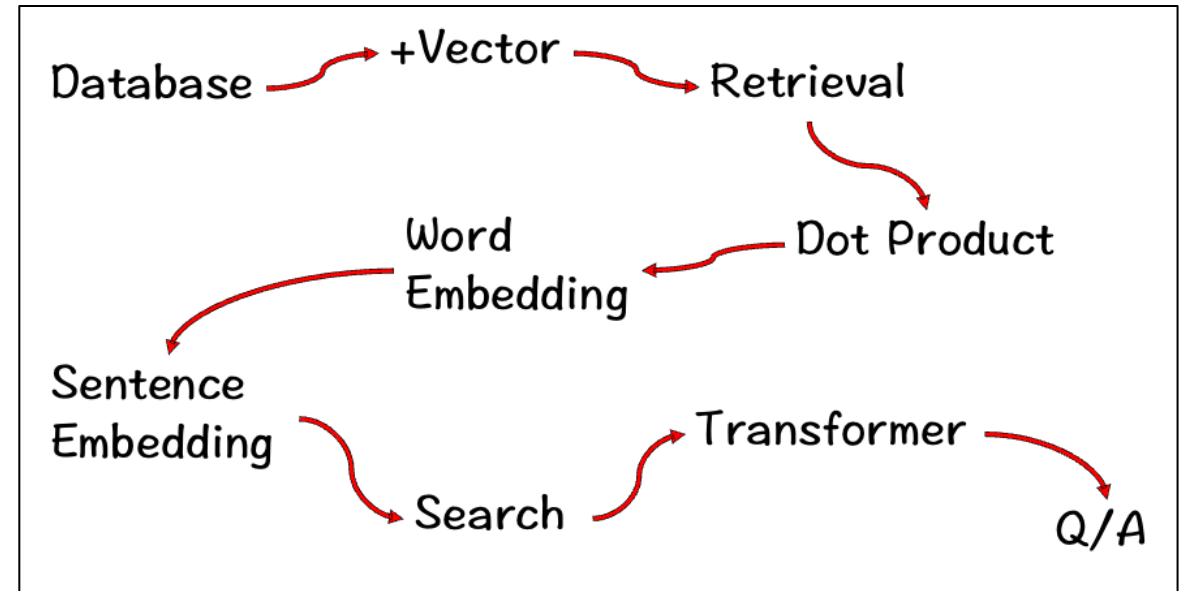
				dog
				2
				1
				0
cat	1	2	0	4

How to compute dot product with multiple vectors?

Example:

	1	1
	2	1
	3	1
2	2	0
6	4	

	dog	bat
	2	0
	1	1
	0	2
cat	1	2
	4	2



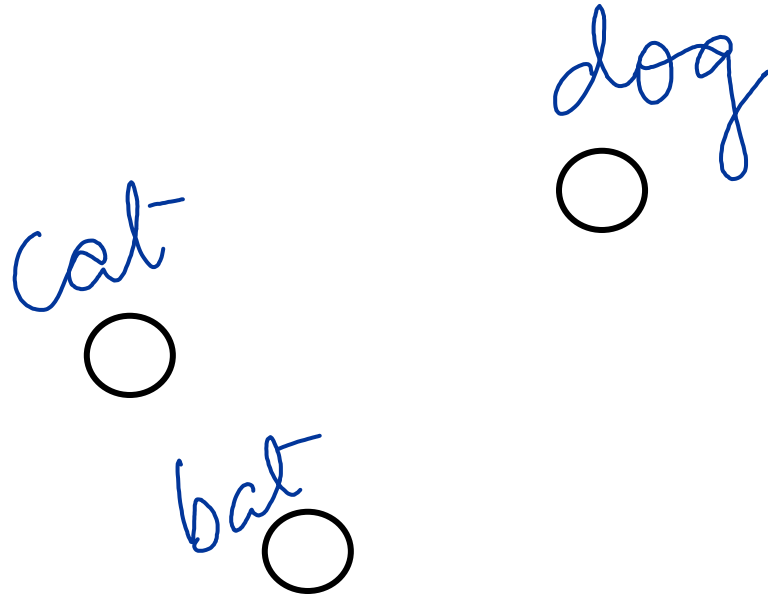
Word Embedding

Beginner's Guide to Vector Databases - AI by Hand 🖋️



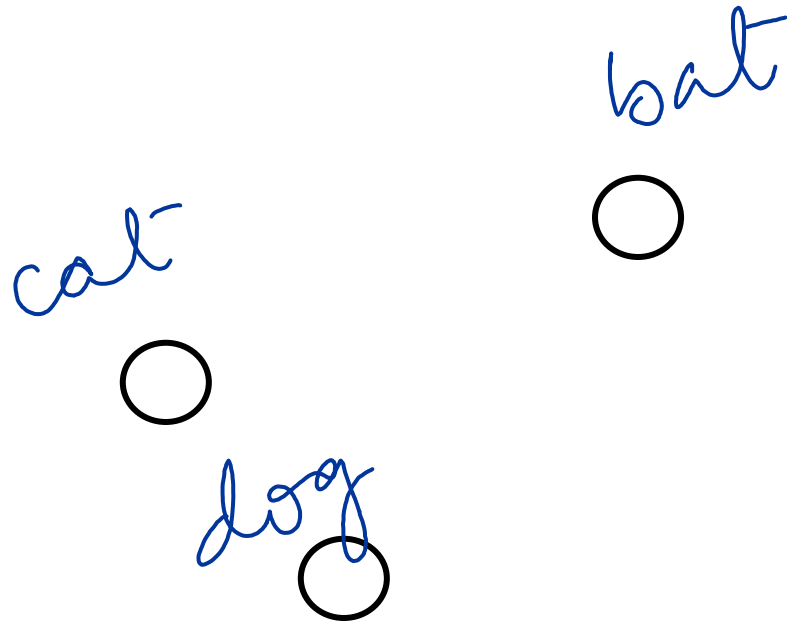
University of Colorado
Boulder

Where are dog, cat and bat in the “name” space?



Where are dog, cat and bat in the “name”
space?

~~name~~
semantic
space



Which embedding is better?

Embedding 1

dog	cat	bat
2	1	0
1	2	1
0	0	2

Embedding 2

dog	cat	bat
2	0	1
1	1	0
0	2	2

Which embedding is better?

Desired
dot
product
similarity

dog	cat	bat
2	1	0
1	2	1
0	0	2

Embedding 1

dog	cat	bat
2	1	0
1	2	1
0	0	2

Embedding 2

dog	cat	bat
2	0	1
1	1	0
0	2	2

dog	2	1	0
cat	1	2	0
bat	0	1	2

	H	L
H		L
L	L	

dog	2	1	0
cat	1	2	0
bat	0	1	2

	4	1
4		2
1	2	

dog	2	1	0
cat	0	1	2
bat	1	0	2

	1	2
1		4
2	4	

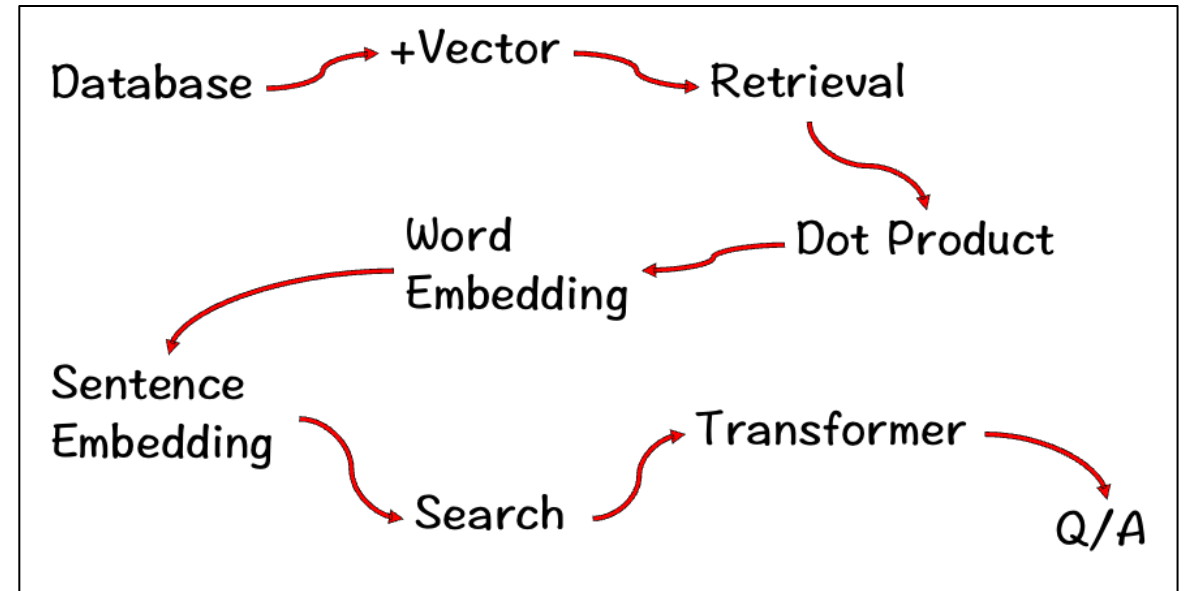
we
want
this

H \Rightarrow high

L \Rightarrow low

3x3 matrix

Yellow is better



Sentence Embedding

Beginner's Guide to Vector Databases - AI by Hand 🖋️




University of Colorado
Boulder

How to embed sentences?

id	comment	user	emb
1	How are you?	John	?
2	Who are you?	Mary	?

“How are you” → word embedding vectors

how are you



a	an	the	how	why	who	what	are	is	am	be	was	you	we	I	they	she	he	she	me	him	her
0	-1	0	1	0	1	0	0	-1	1	0	0	0	3	1	0	-1	0	0	0	-1	0
2	0	2	0	0	0	-1	1	0	0	0	2	1	0	2	0	2	0	0	2	0	0
-1	0	-1	1	2	0	0	1	0	1	-1	0	0	-1	0	3	0	0	-1	0	2	-1
0	1	0	0	1	0	1	0	1	0	1	-2	0	0	0	1	0	1	0	1	0	1

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

Word vectors → Sentence vector

Method 1: Concatenate

how	are	you
-----	-----	-----

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

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

→ This might not
be scalable

Word vectors \rightarrow Sentence vector

Method 2: Average

how	are	you
1	0	0
0	1	1
1	1	0
0	0	0

$\frac{1}{3}$
$\frac{2}{3}$
$\frac{2}{3}$
$\frac{0}{3}$

Vector(4)

id	comment	user	emb
1	How are you?	John	$[\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{0}{3}]$
2	Who are you?	Mary	

“Who are you” → word embedding vectors

who	are	you
-----	-----	-----

a	an	the	how	why	who	what	are	is	am	be	was	you	we	I	they	she	he	she	me	him	her
0	-1	0	1	0	1	0	0	-1	1	0	0	0	3	1	0	-1	0	0	0	-1	0
2	0	2	0	0	0	-1	1	0	0	0	2	1	0	2	0	2	0	0	2	0	0
-1	0	-1	1	2	0	0	1	0	1	-1	0	0	-1	0	3	0	0	-1	0	2	-1
0	1	0	0	1	0	1	0	1	0	1	-2	0	0	0	1	0	1	0	1	0	1

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

Word vectors \rightarrow Sentence vector

Method 2: Average

who	are	you
1	0	0
0	1	1
0	1	0
0	0	0

$\frac{1}{3}$
$\frac{2}{3}$
$\frac{1}{3}$
$\frac{0}{3}$

id	comment	user	emb
1	How are you?	John	$[\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 0]$
2	Who are you?	Mary	$[\frac{1}{3}, \frac{2}{3}, \frac{1}{3}, 0]$

How to query by SQL?

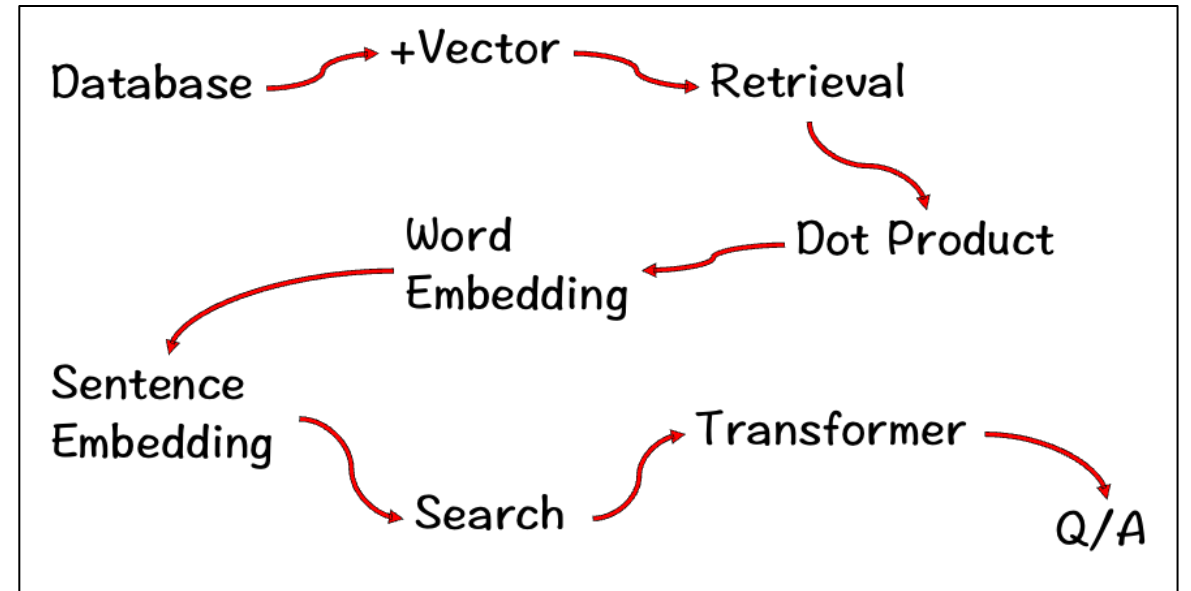
dot product / similarity

```
SELECT comment, emb<*>[2/3, 1/3, 3/3, 4/3] AS score  
FROM posts  
ORDER BY score ASC | DESC;
```


How to query using a high-level API?

```
query = Query(post_index)
    .find(post)
    .similarity(relevance_space.text, Param("comment"))
app.query(query, comment = "who are you?" )
```

Source: Superlinked.com 



Search

Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

K-Nearest Neighbor, K=3, Dot-Product

Database

ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
emb																				

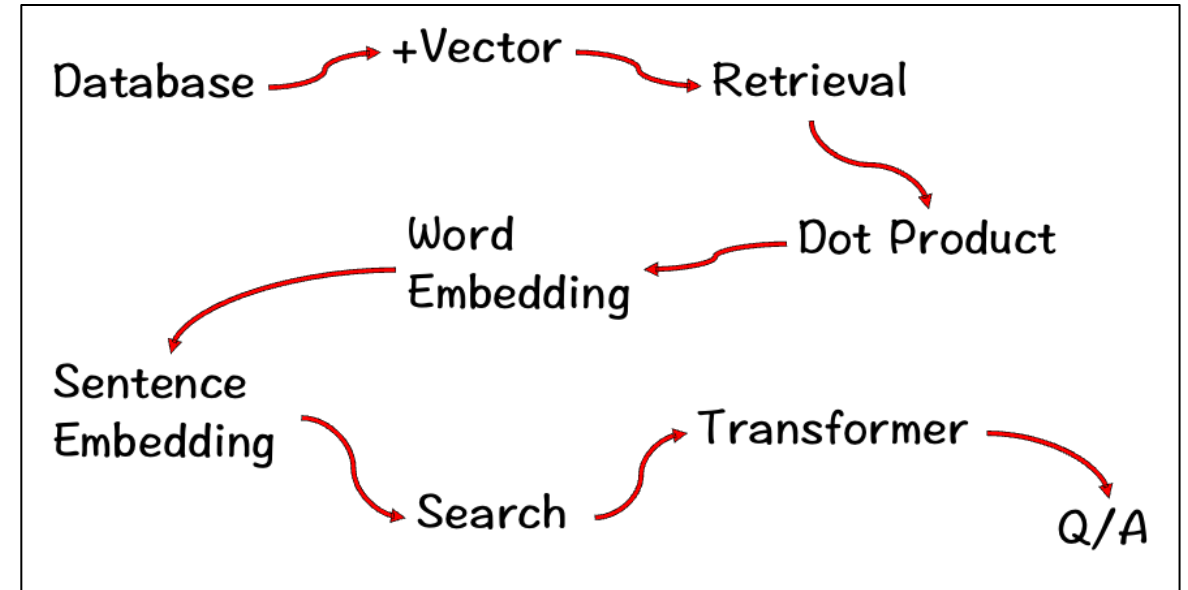
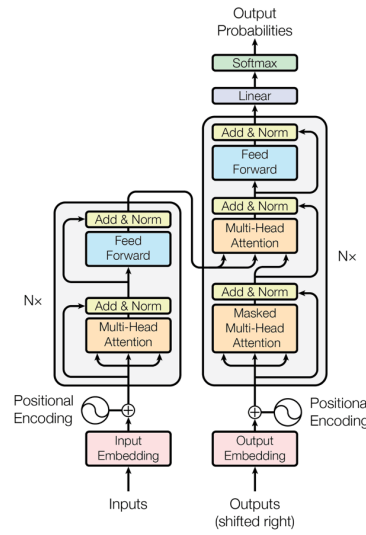
Query					9	-8	9	9	0	3	1	-6	0	11	3	13	-2	6	15	-9	7	6	-5	8
{ max min }														3		2			1					

K-Nearest Neighbor, $K=3$, L2 \rightarrow Euclidean Dist (-)

Database	
ID	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
emb	
Query	
	6 8 9 9 1 10 0 9 12 15 2 13 12 6 15 9 7 6 5 8
{ max min }	
	2 1 3

* 20m records \Rightarrow this process will be slow

* ANN \Rightarrow retrieval is faster



Transformer

Beginner's Guide to Vector Databases - AI by Hand 🖋️

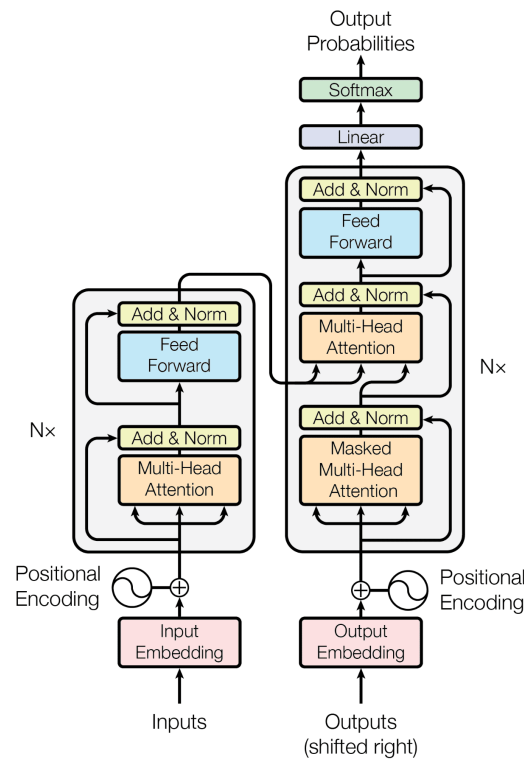


University of Colorado
Boulder

How to use a Transformer to get a sentence embedding vector?

Word
Embedding
Vectors

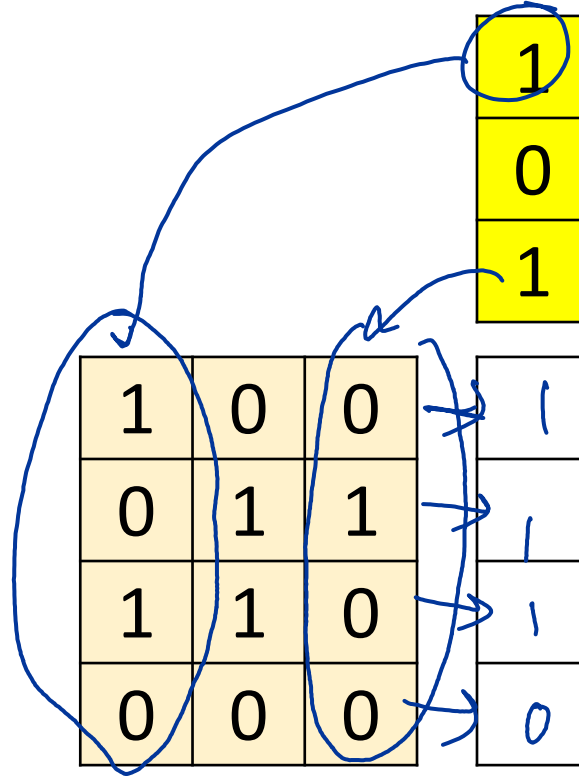
1	0	0
0	1	1
1	1	0
0	0	0



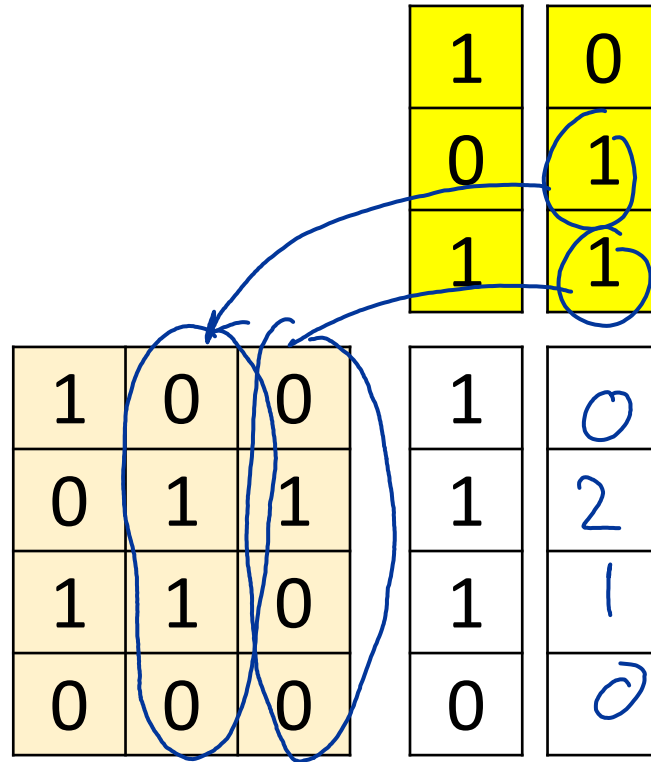
Sentence
Embedding
Vector



How to combine across positions?



How to combine across positions?

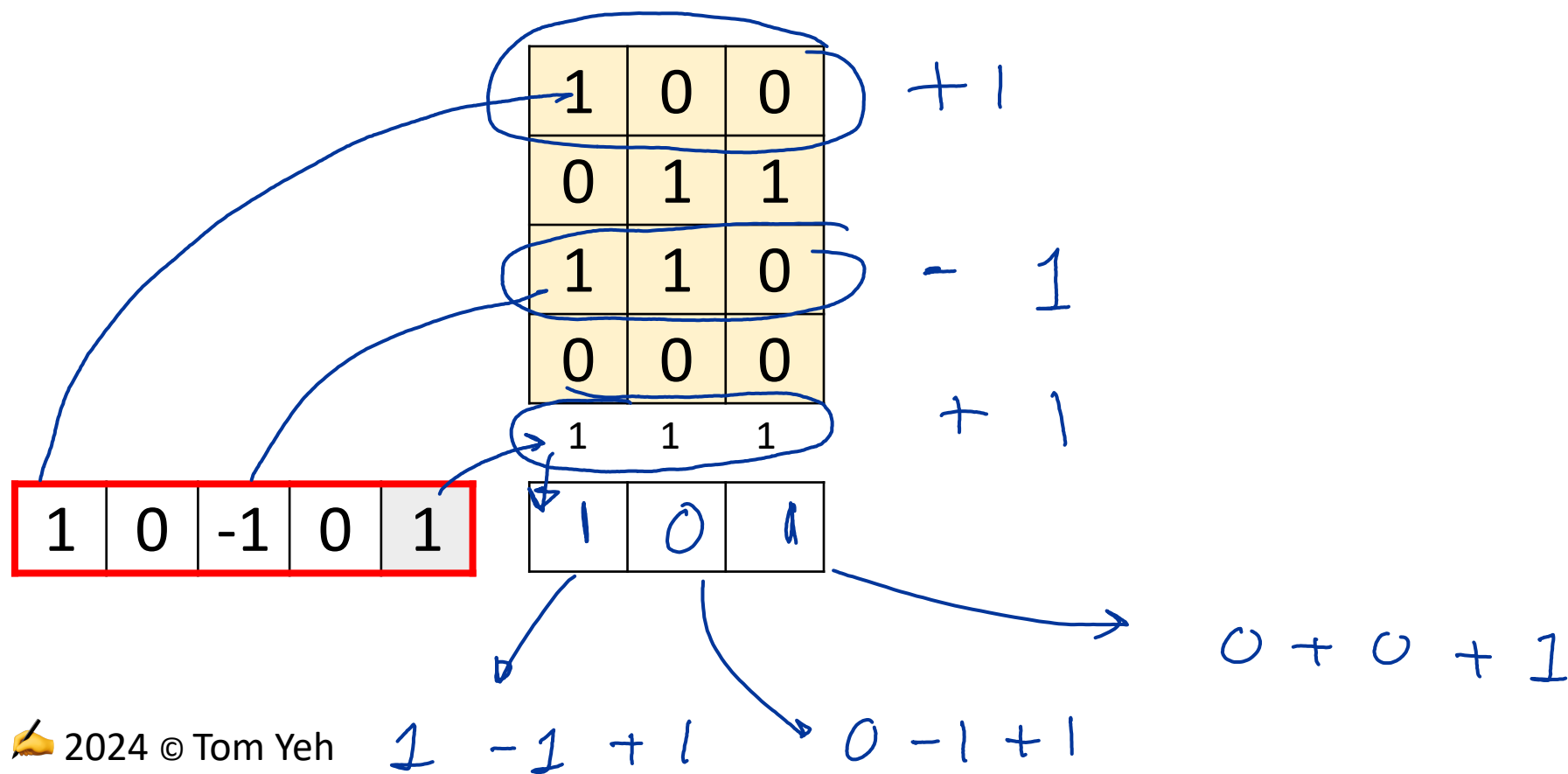


How to combine across positions?

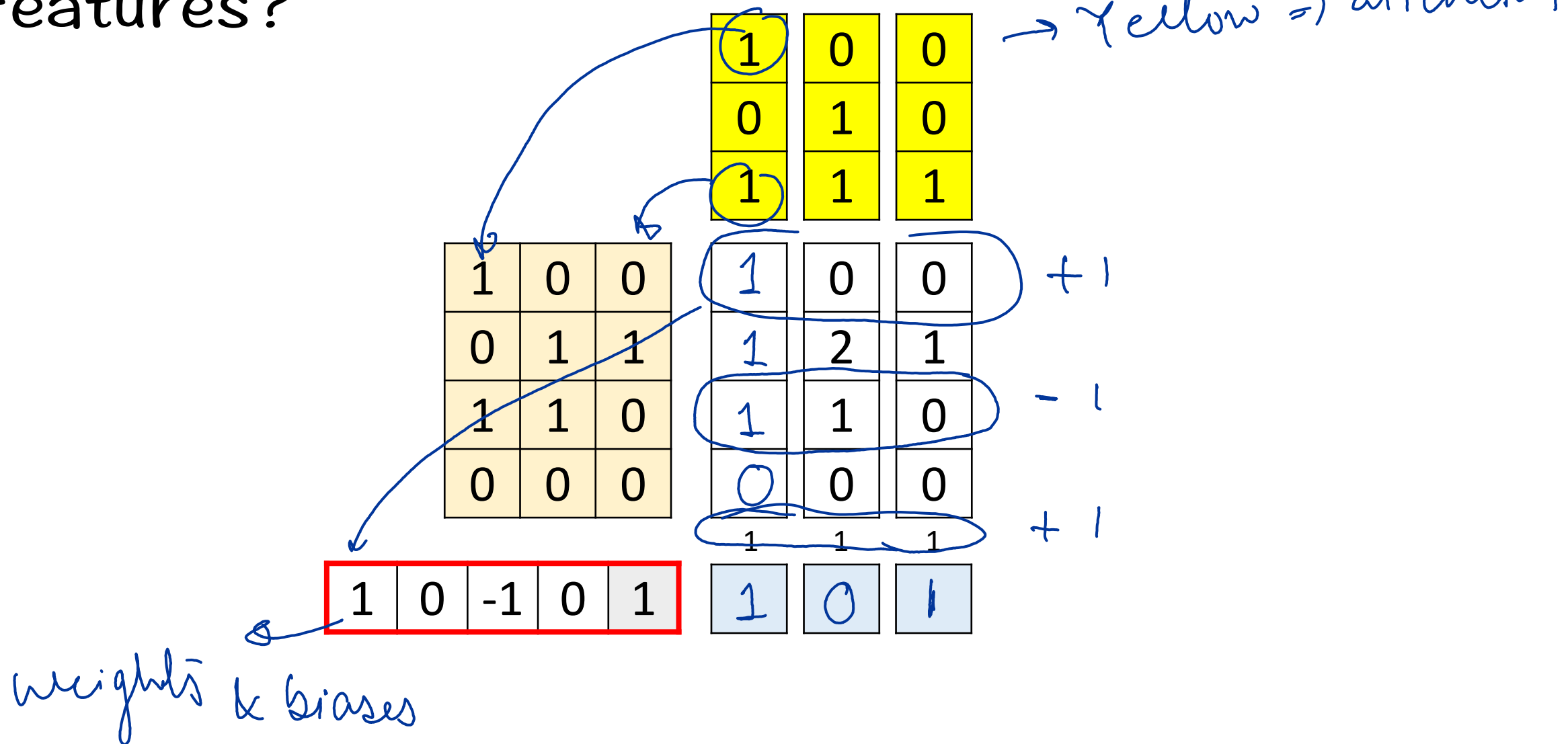
1	0	0
0	1	1
1	1	0
0	0	0

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

How to combine across features?



How to combine across positions and features?



How to use a Transformer to get a sentence embedding vector?

Word
Embedding
Vectors

1	0	0
0	1	0
1	1	1

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

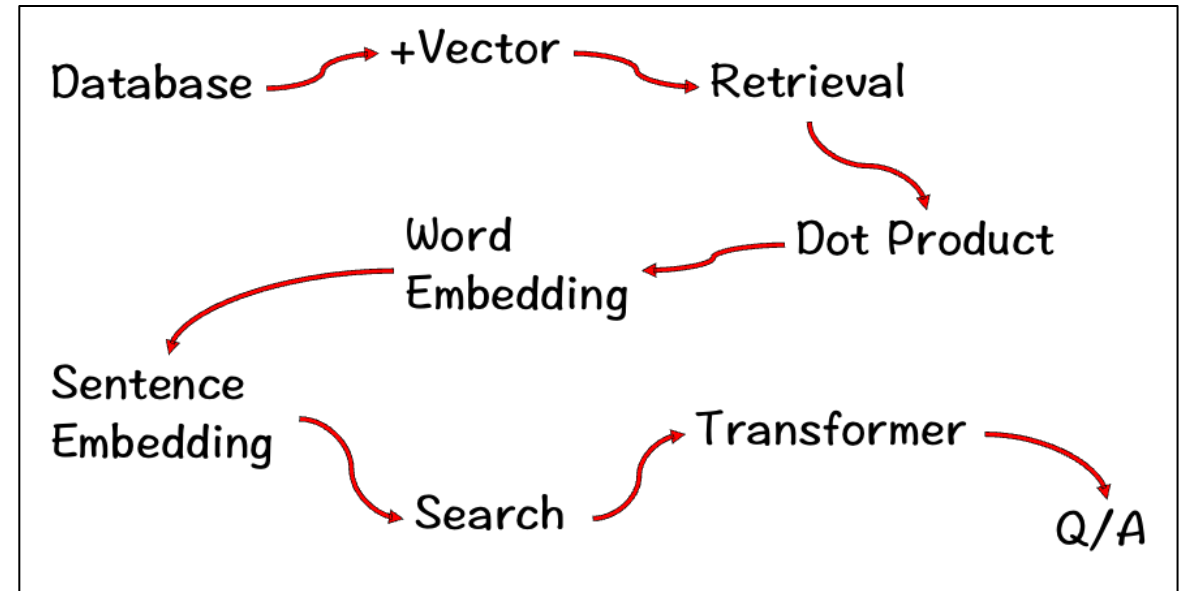
1	0	0
1	2	1
1	1	0
0	0	0

Sentence
Embedding
Vector

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

1	1	1
1	0	1
2	3	1
1	1	1
1	1	0

2/3
6/3
3/3
2/3



Q/A

Beginner's Guide to Vector Databases - AI by Hand 🖋️



University of Colorado
Boulder

