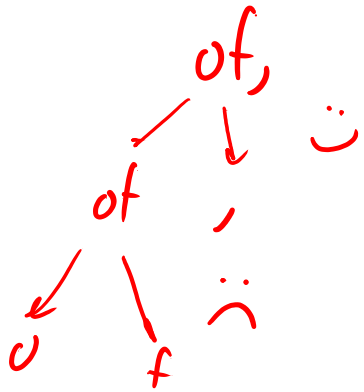
LLM

Large Language Models

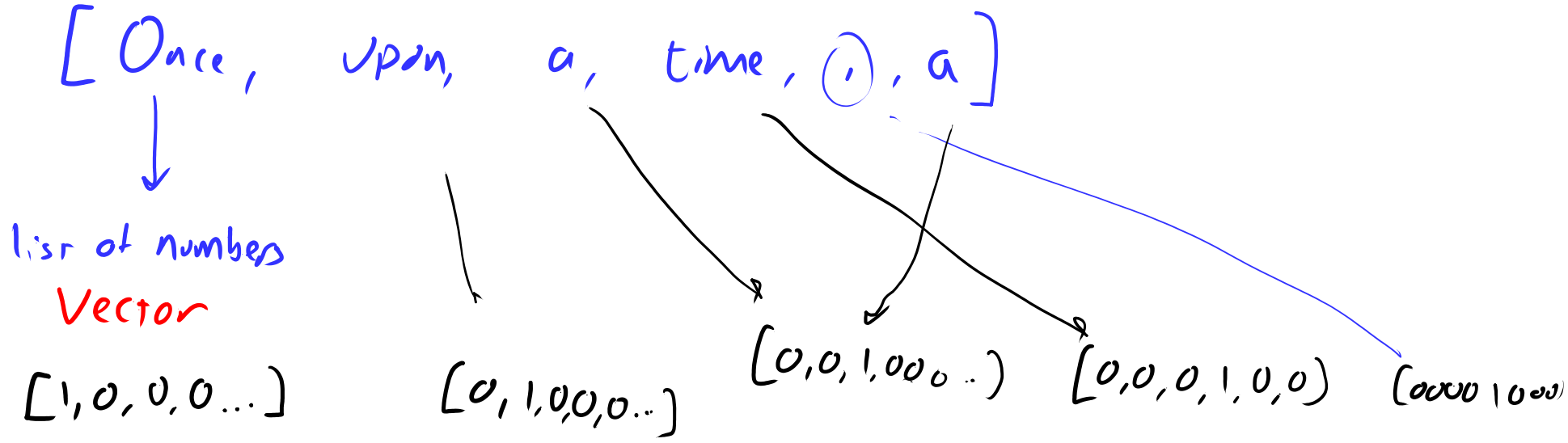
text → bytes → Unicode

Lex
Lexical Analysis
tokens
↳ each should be 1 meaning
(atomic)

Of of of, :)



(sound) ← hard



One-hot encoding

Vector : list of numbers $(3, 7, -2, 0.53)$

inner product:

$$[x, y, z) \quad [a, b, c] \longrightarrow xa + yb + zc$$

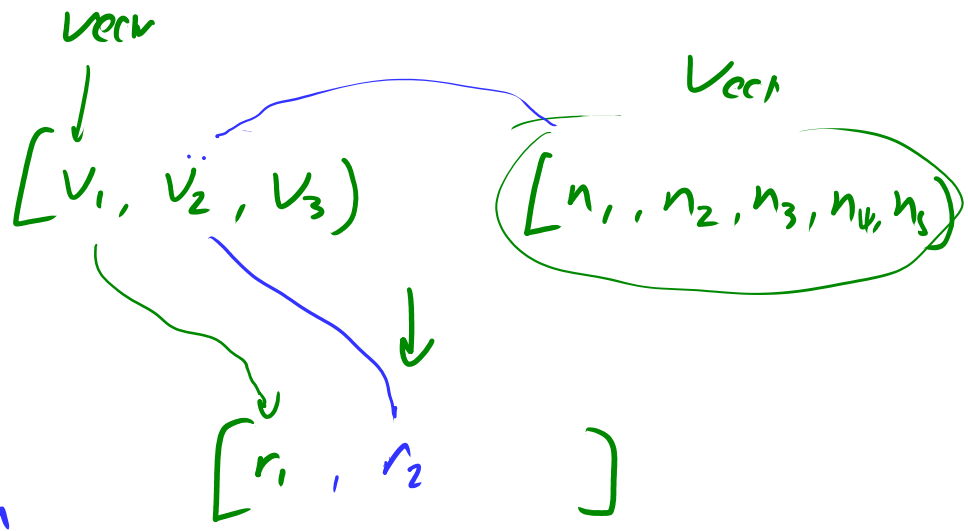
multiplies pairs of num
adds them

input: 1 vector

linear operator: list of vectors

output: 1 vector

linear algebra



Word similarity

likelihood of being swapped

of

king

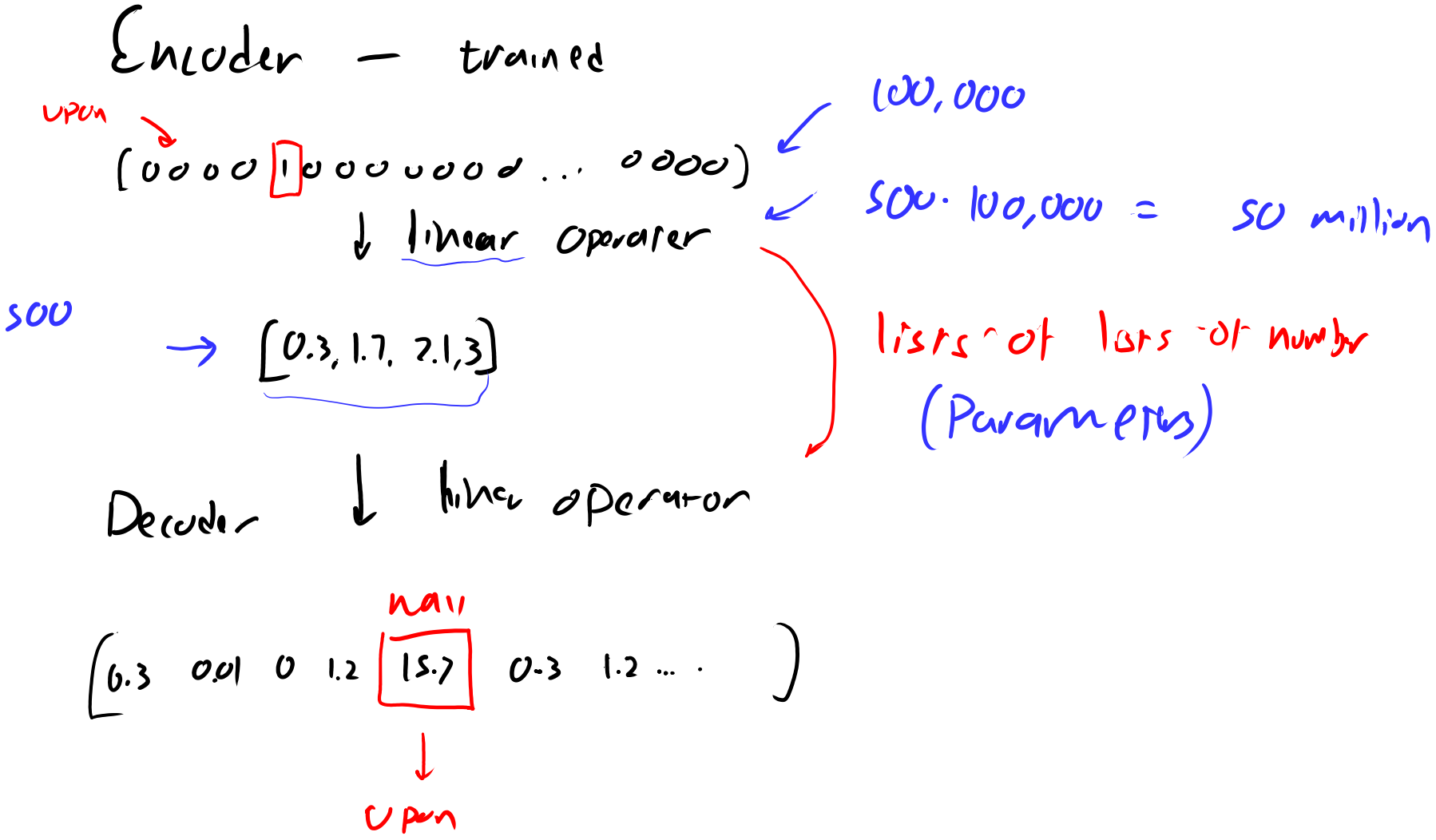
queen

one of the

the king sat | the queen sat

said Queen Elizabeth

⋮


Encoder — trained
upon
(0000 1 000 0000 ... 0000)
100,000
500 · 100,000 = 50 million
linear operator
500
[0.3, 1.7, 2.1, 3]
lists of lists of number
(Parameters)
Decoder
linear operator
nail
(0.3 0.01 0 1.2 15.7 0.3 1.2 ... .)
upon

Concatenate: stick one after another

$[1,2]$ $(7,8)$ $\rightarrow$ $(1,2,7,8)$

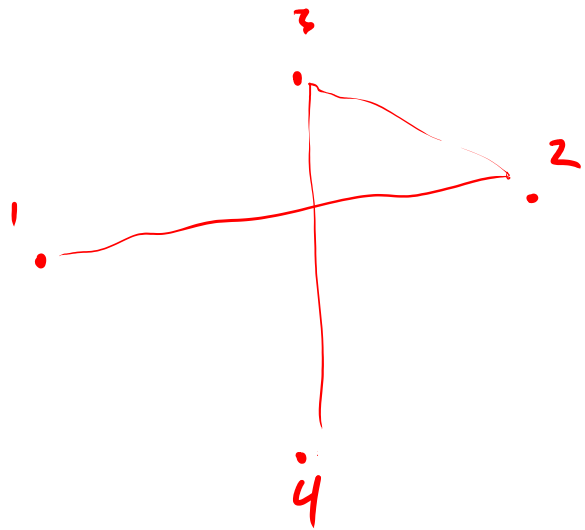
Earth – planet

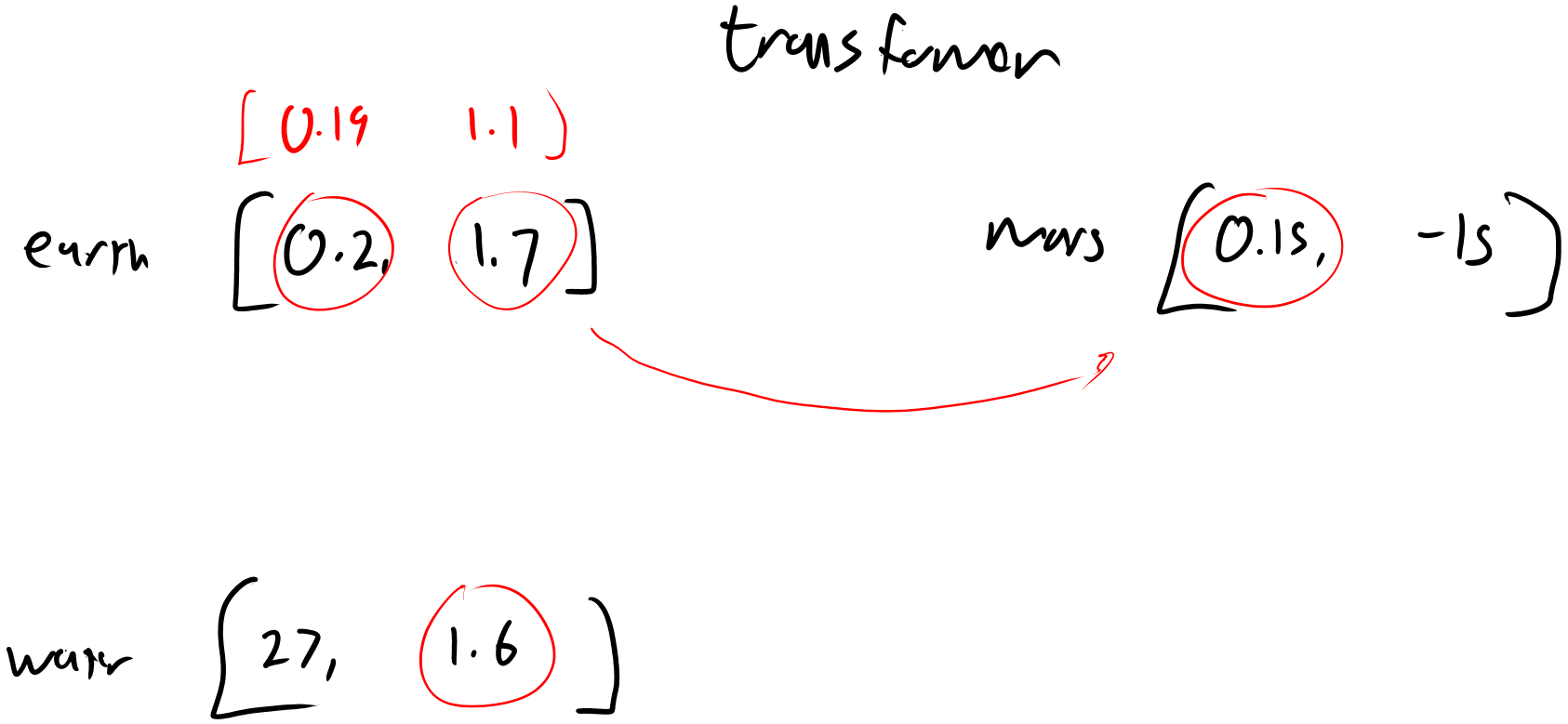
rock

Where is the earth?

I was buried beneath ton of earth

Connect the dots




transformer
[0.19 1.1)
earth [0.2, 1.7]
mars [0.15, -15]
water [27, 1.6]

Artificial Neural Network (ANN)

(Neural net)

linear operation, then rounding

$< \frac{1}{2} \rightarrow 0$

$> \frac{1}{2} \rightarrow 1$