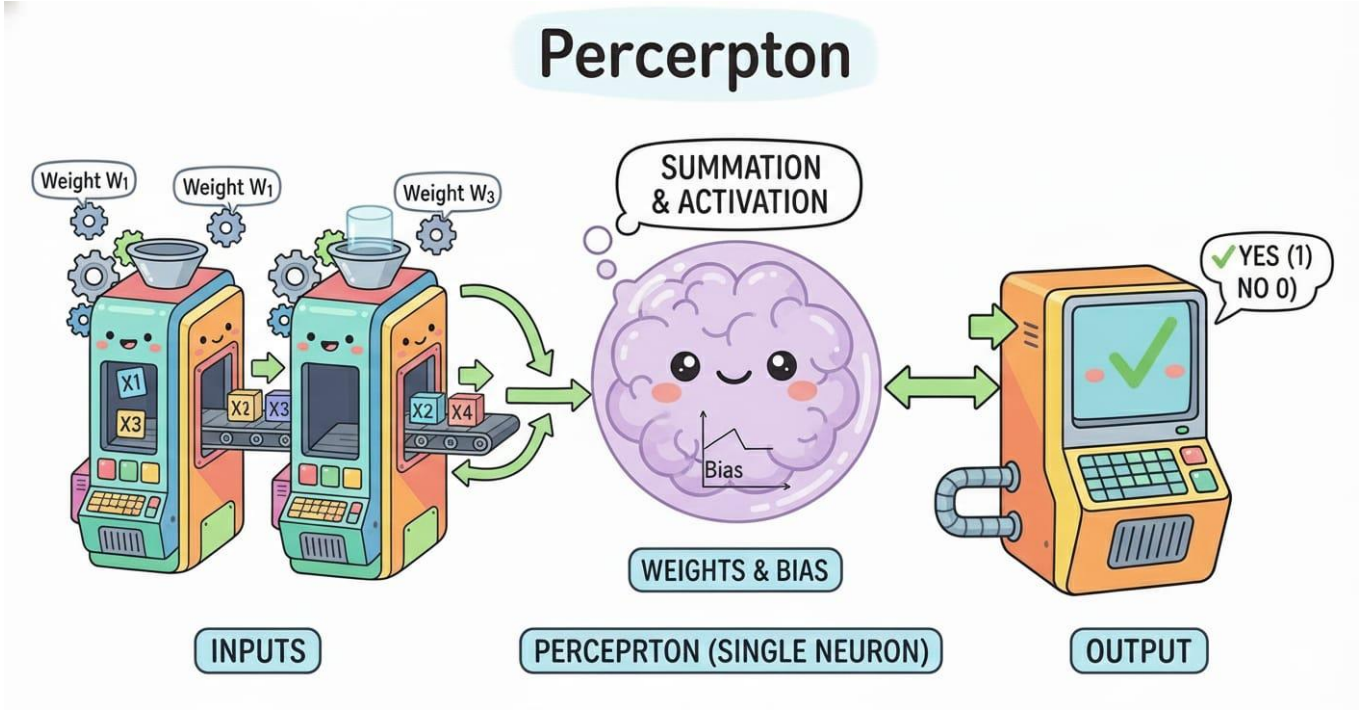
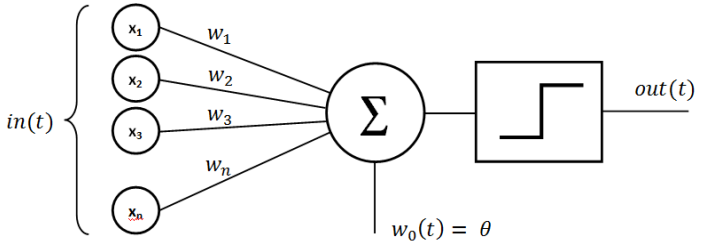
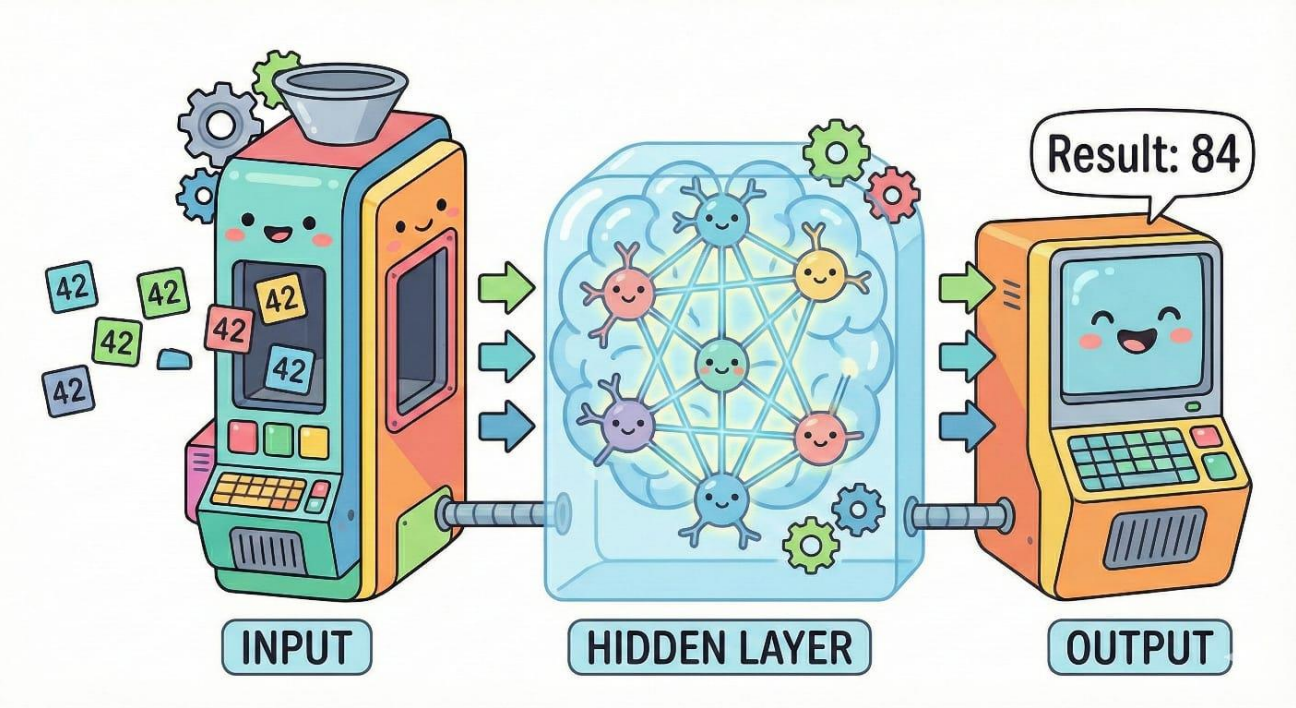
Un breve camino (lúdico) por las redes neuronales

Erick Merino

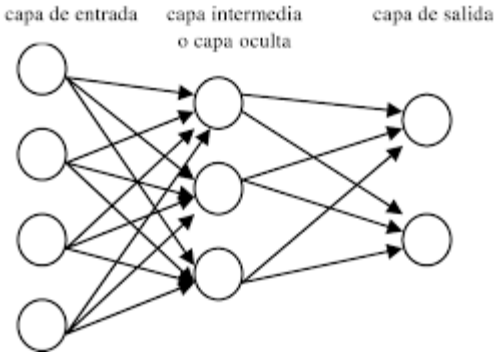


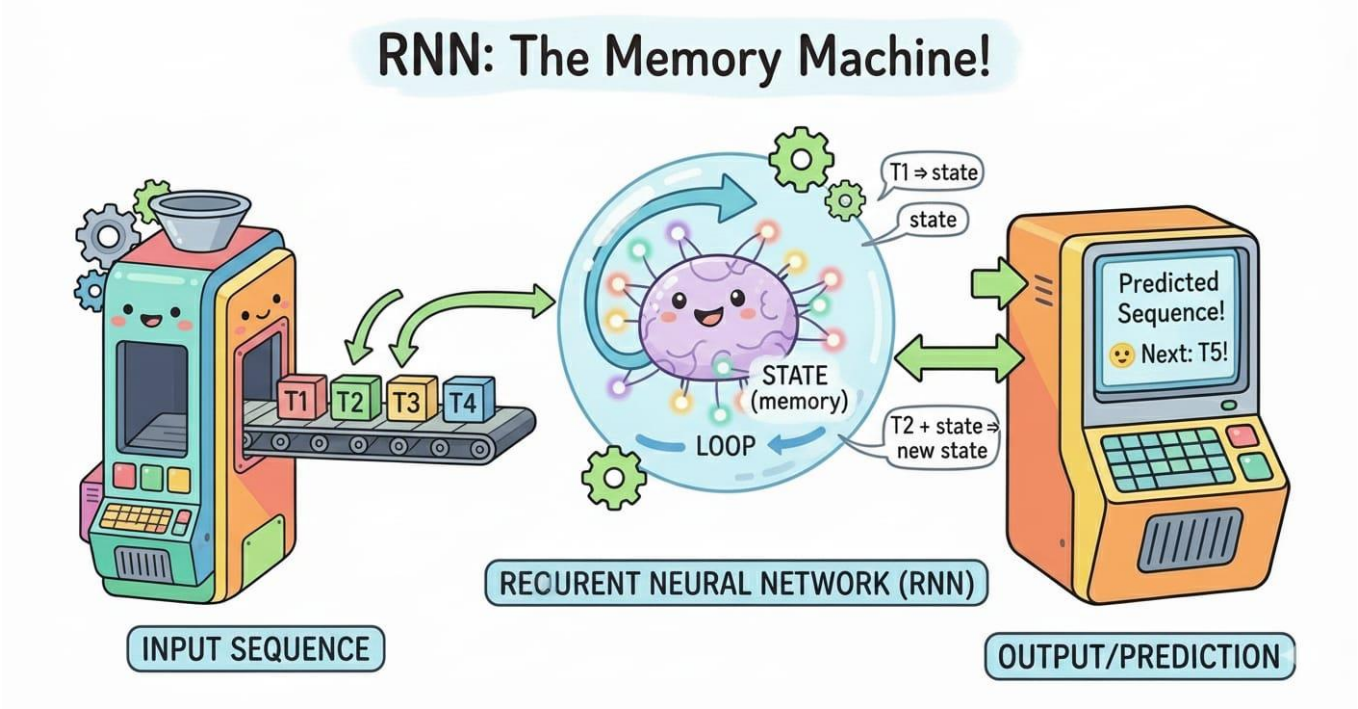
Todo empezó con una neurona



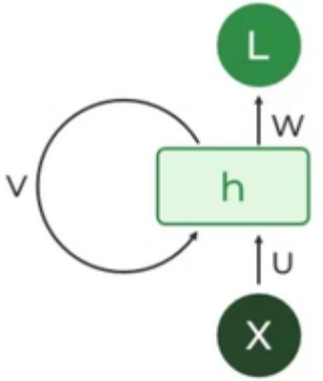


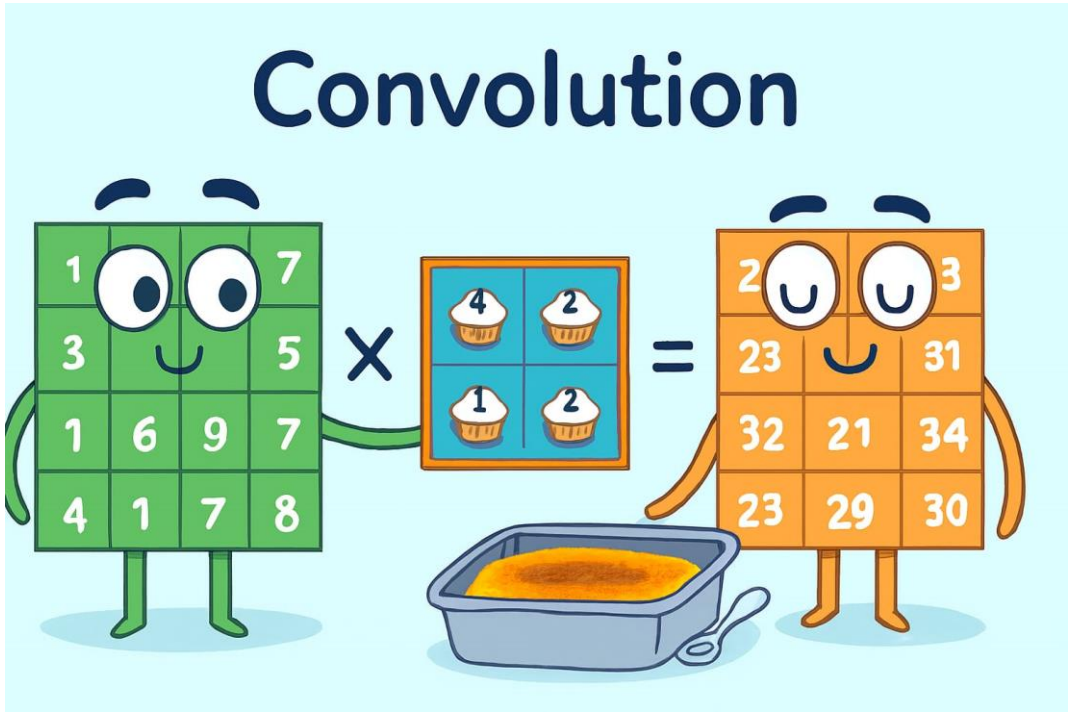
A la cual se le sumaron más



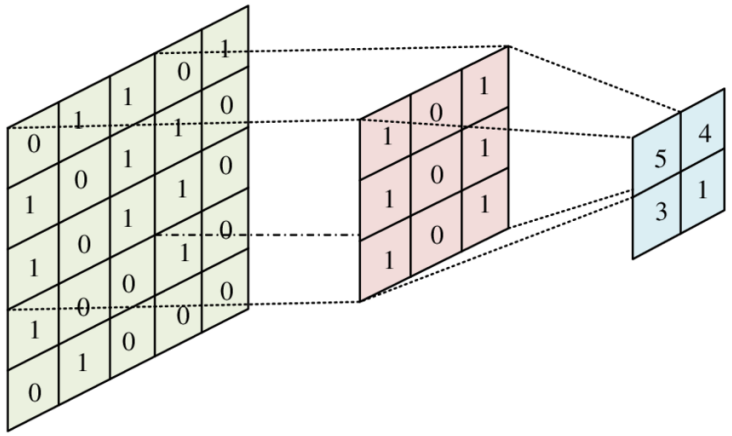


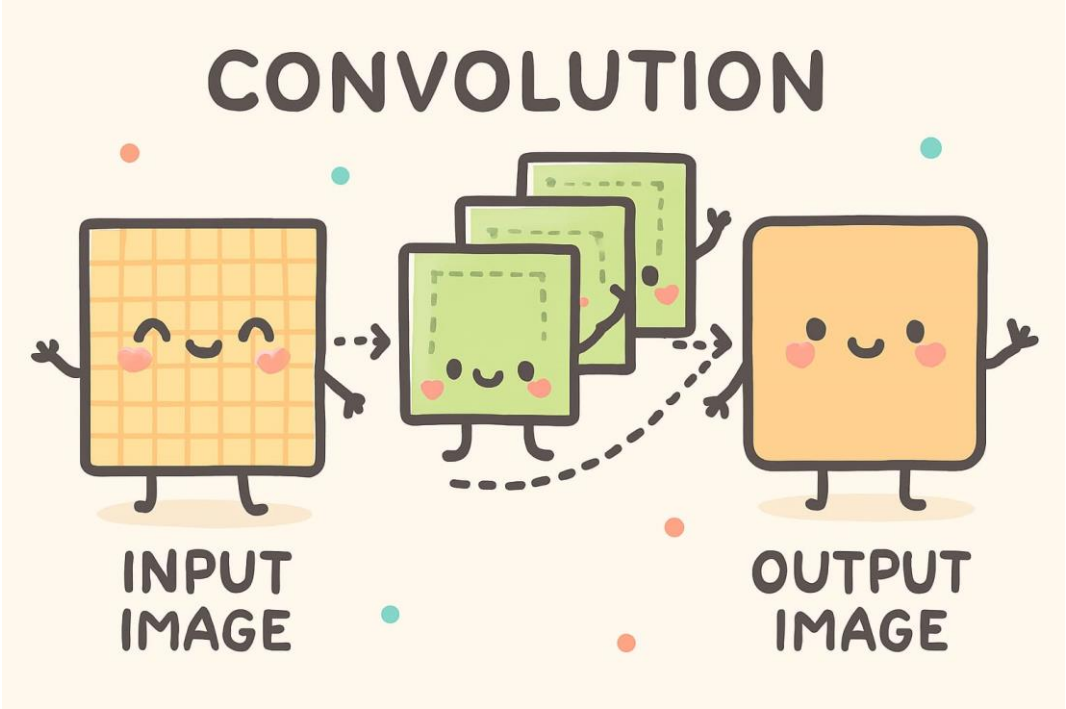
Se les dio memoria



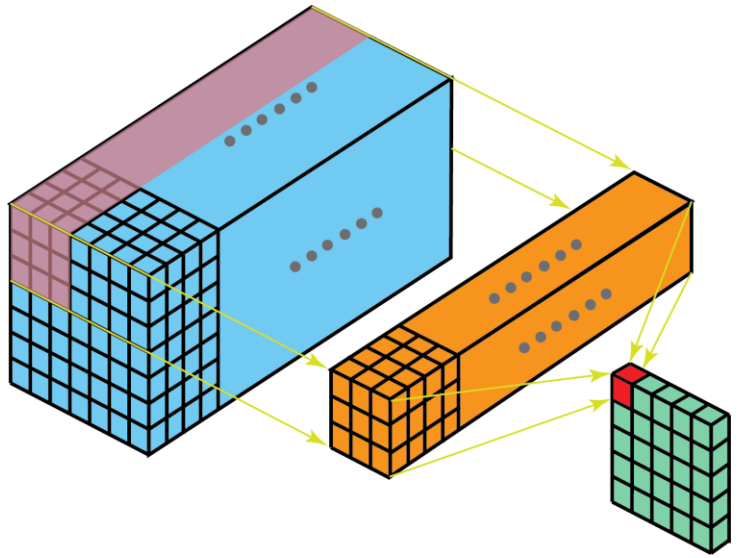


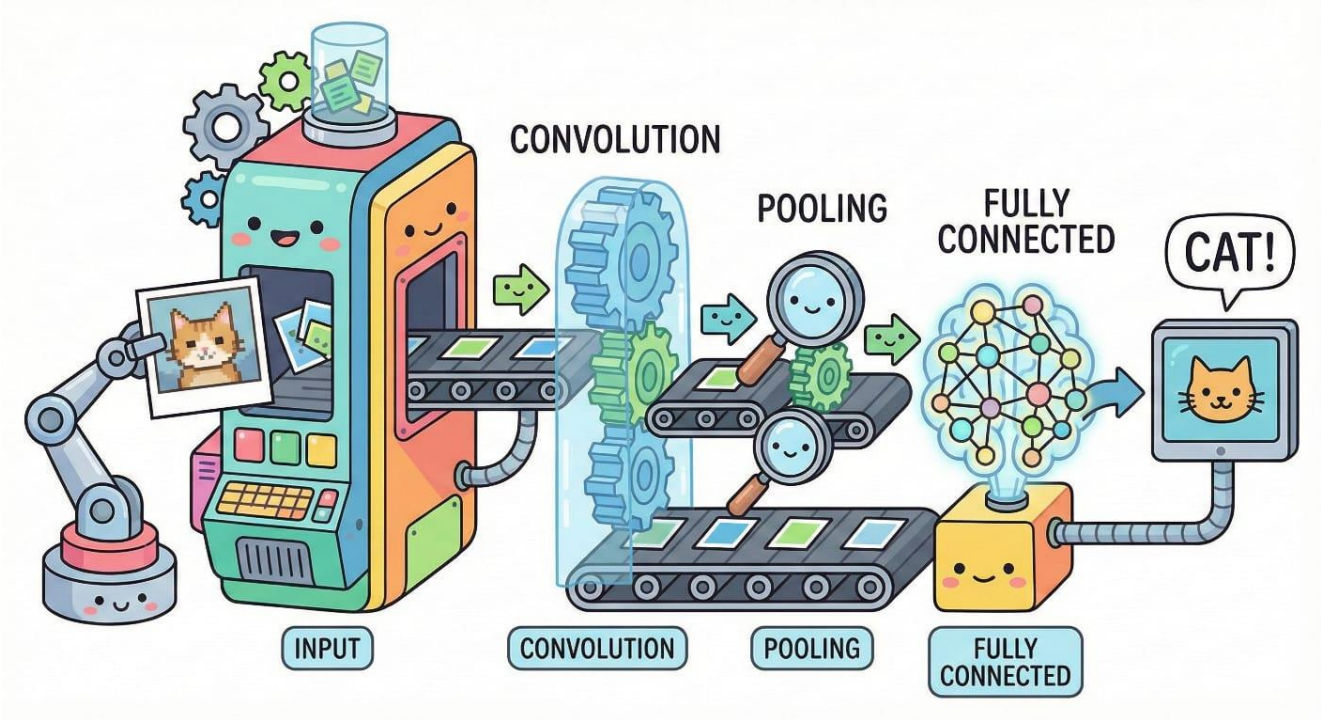
Se las ordenó y multiplicó



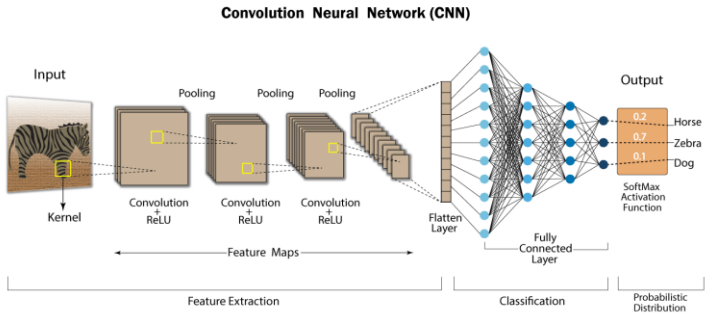


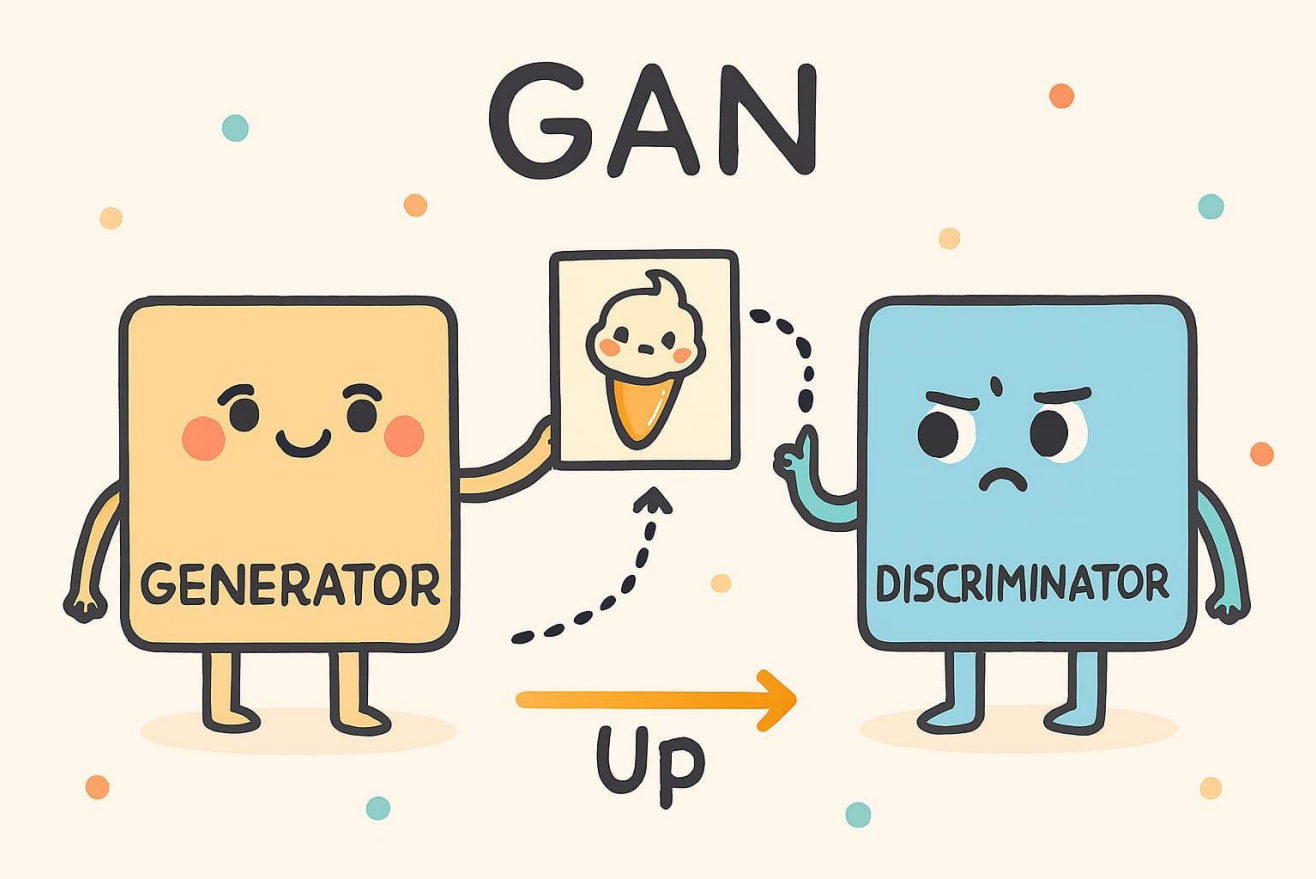
Múltiples veces



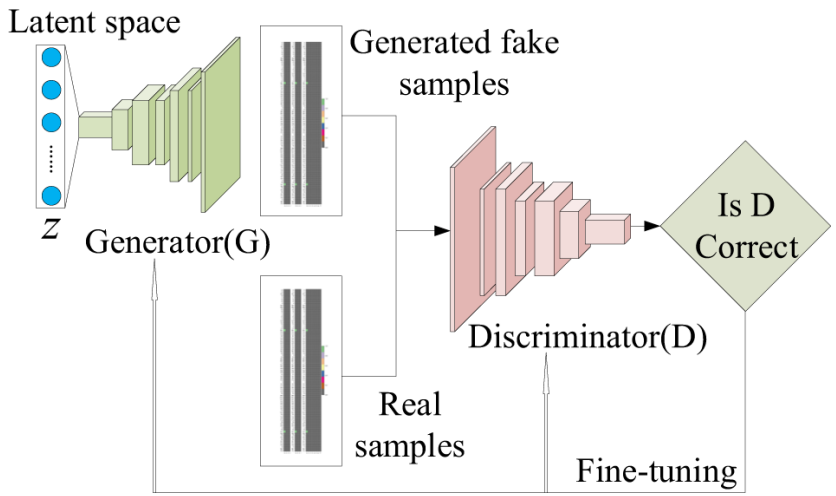


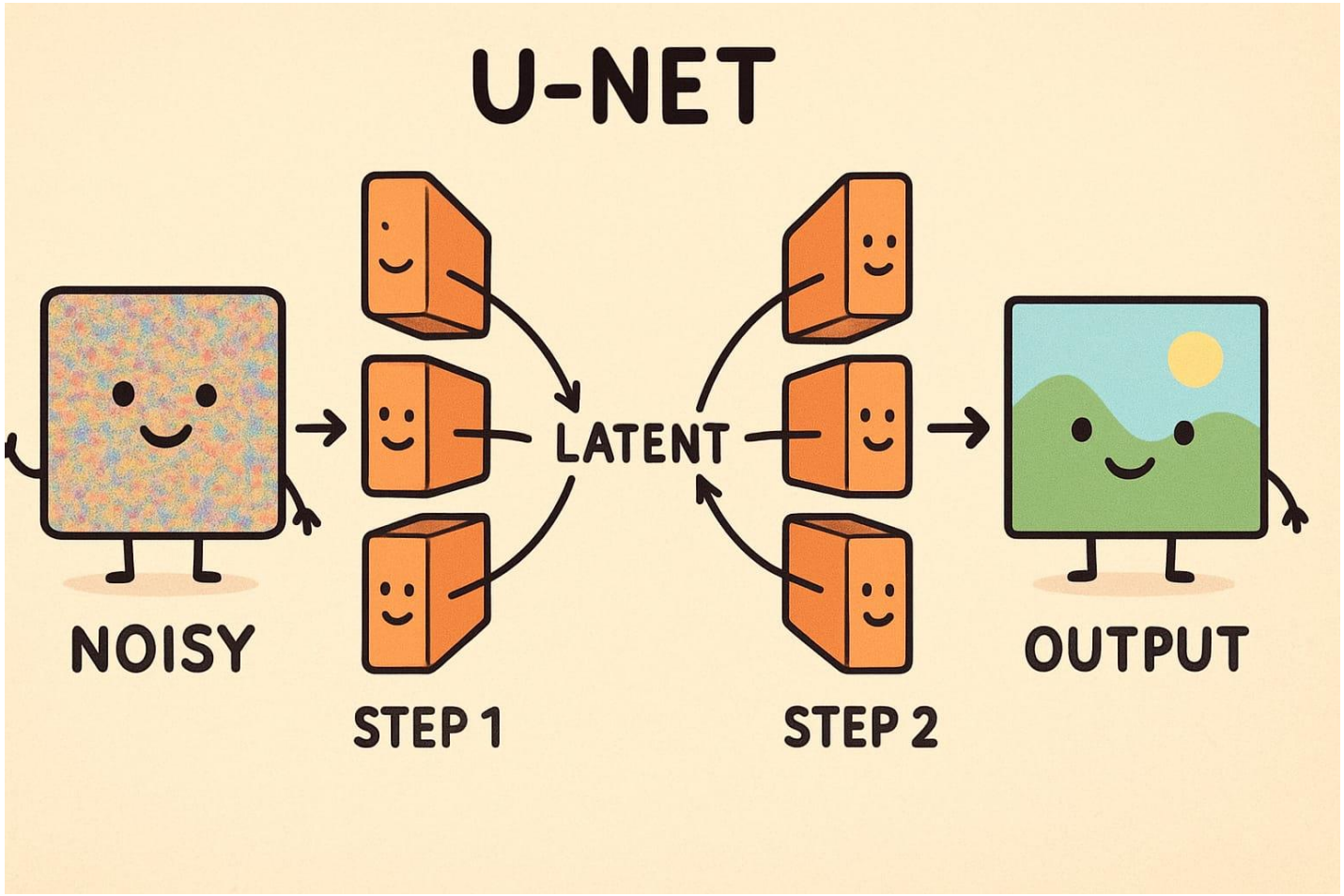
Se crearon complejas máquinas de neuronas



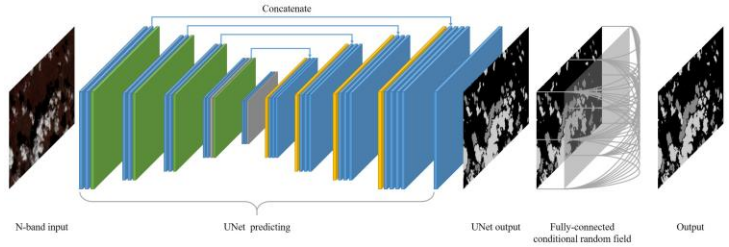


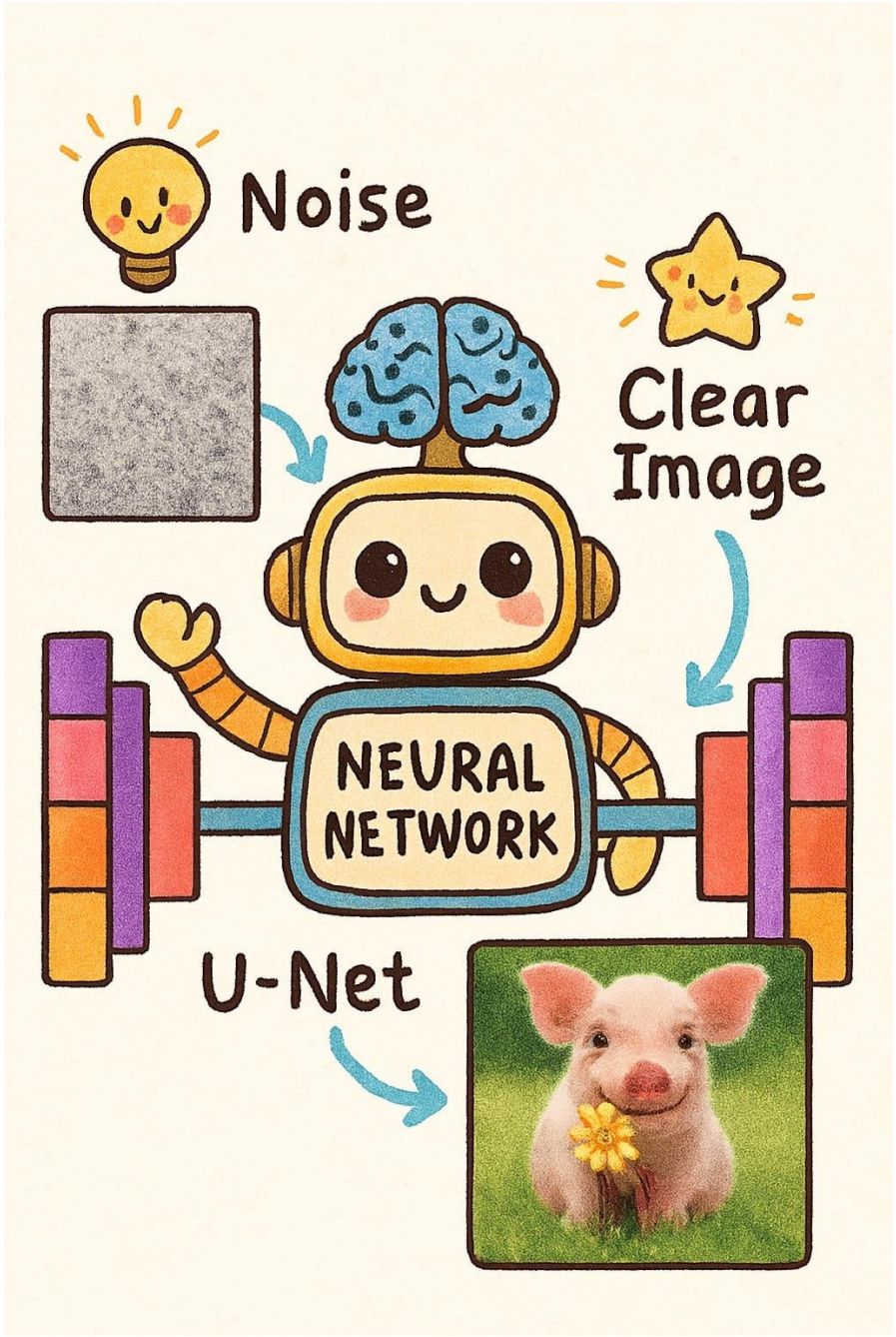
Se las hizo competir



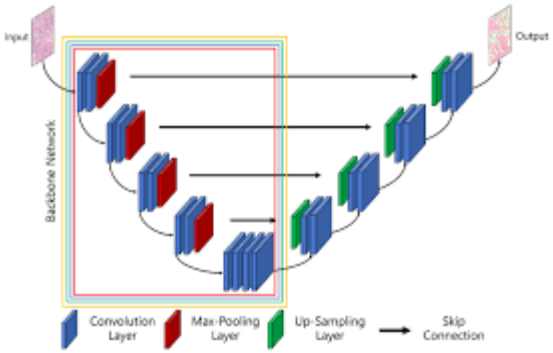


Se las empequeñeció y agrandó

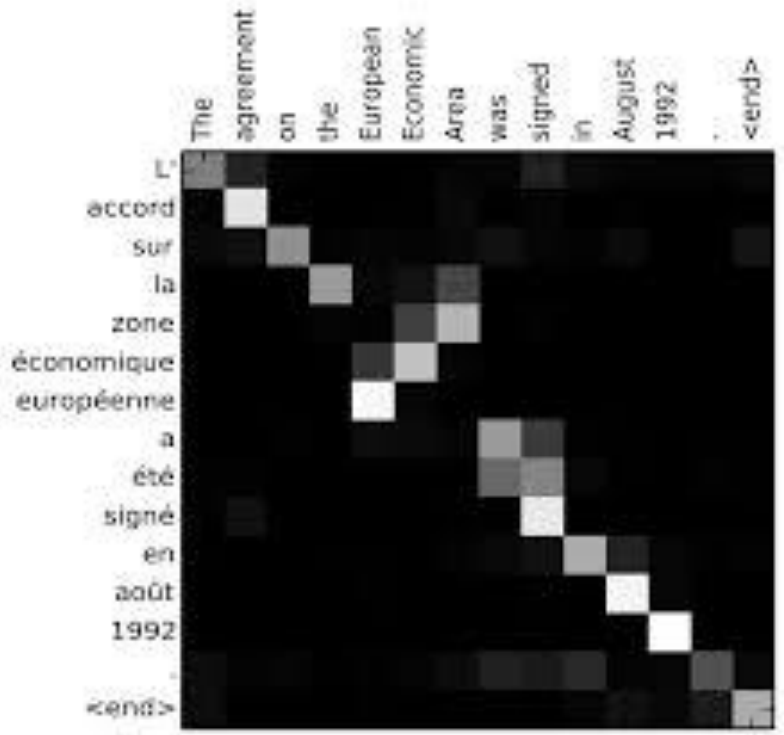


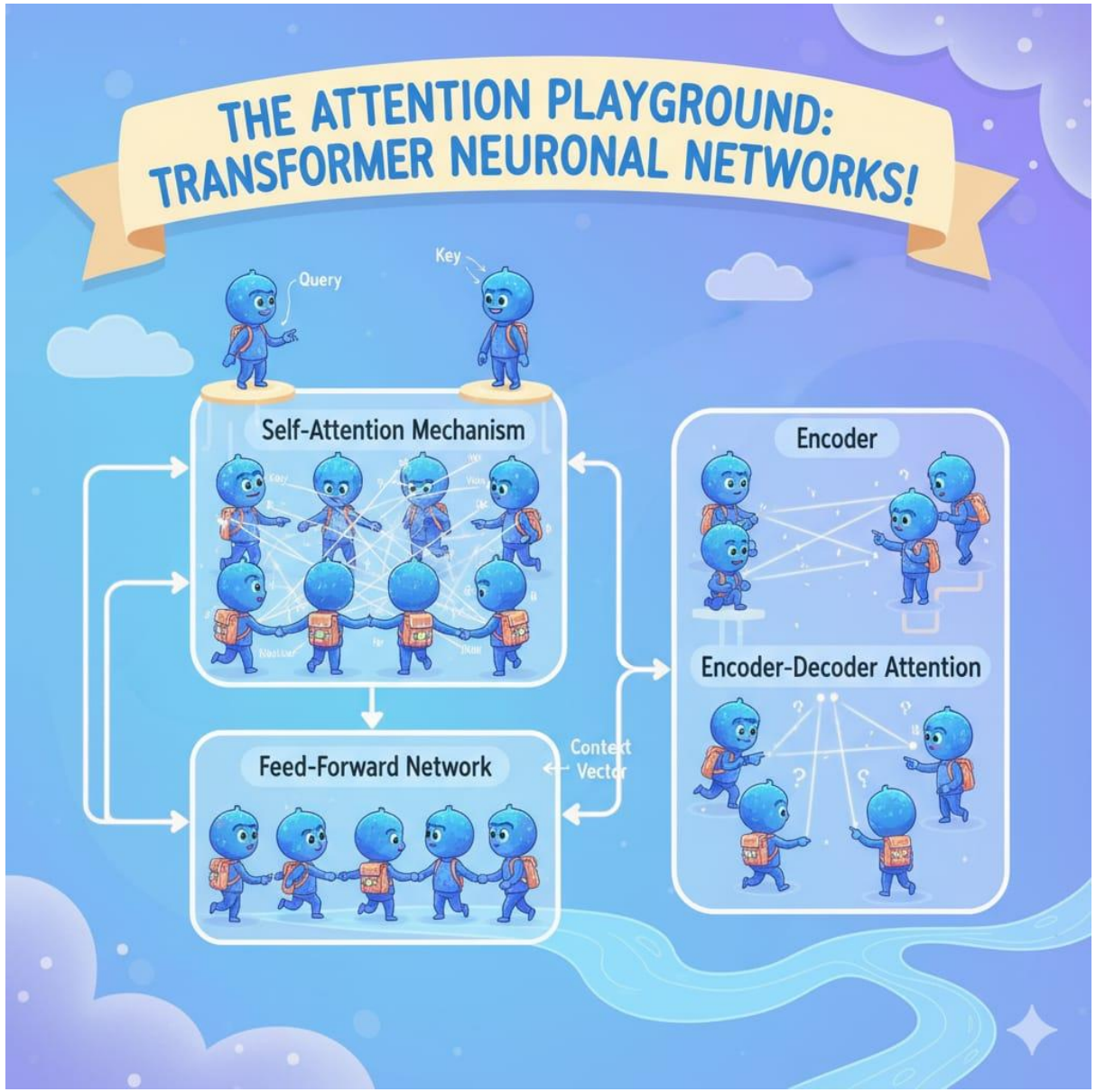


Fueron capaces de generar imágenes

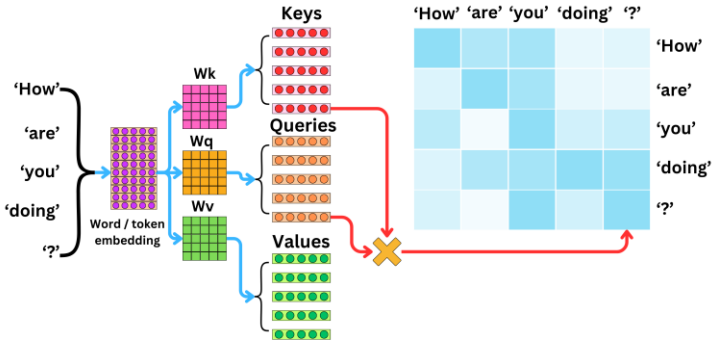


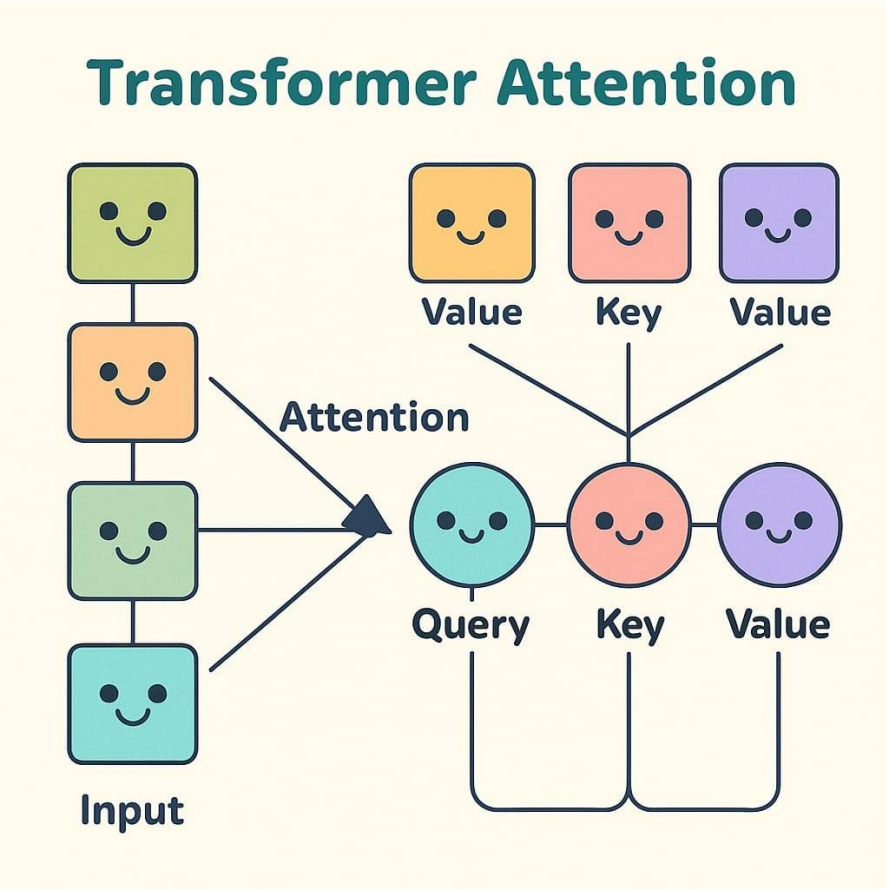
Se las hizo interactuar entre ellas



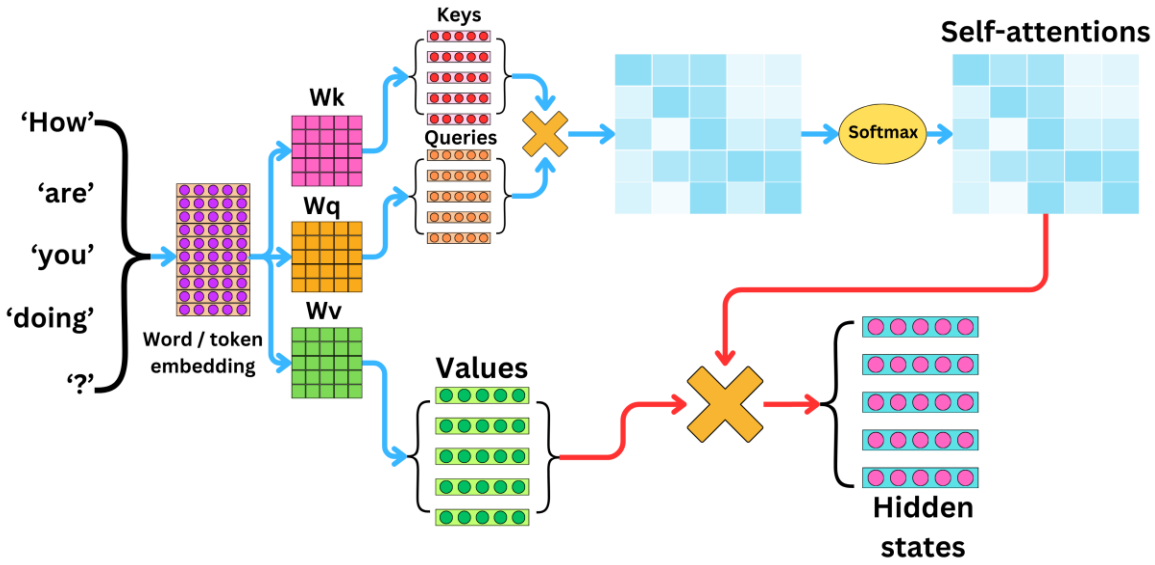

THE ATTENTION PLAYGROUND:
TRANSFORMER NEURONAL NETWORKS!
Query
Key
Self-Attention Mechanism
Encoder
Encoder-Decoder Attention
Feed-Forward Network
Context Vector

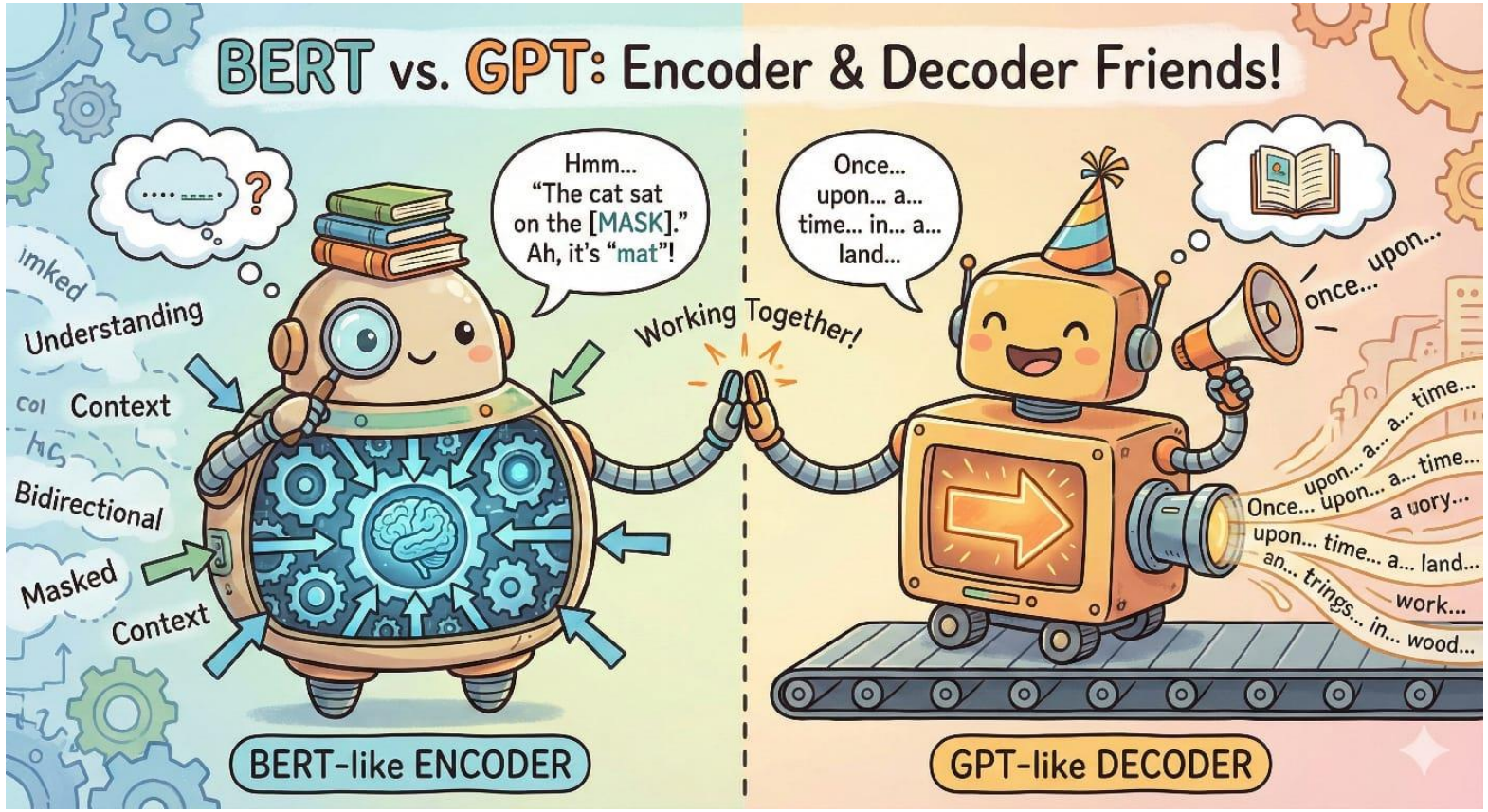
Mirarse


'How'
'are'
'you'
'doing'
'?'
Word / token embedding
Wk
Wq
Wv
Keys
Queries
Values
'How' 'are' 'you' 'doing' '?'
'How'
'are'
'you'
'doing'
'?'

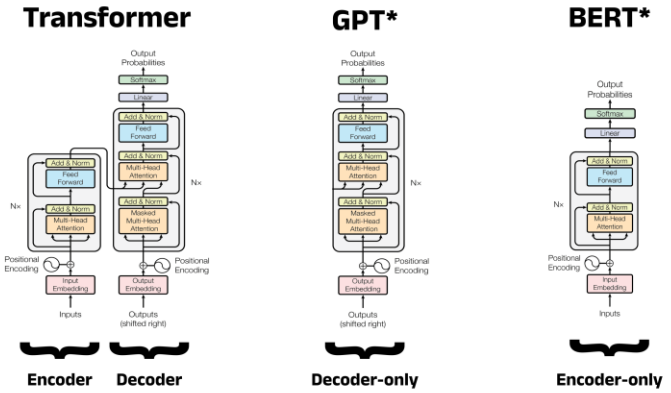


A través de complejos mecanismos, compararse a sí mismas

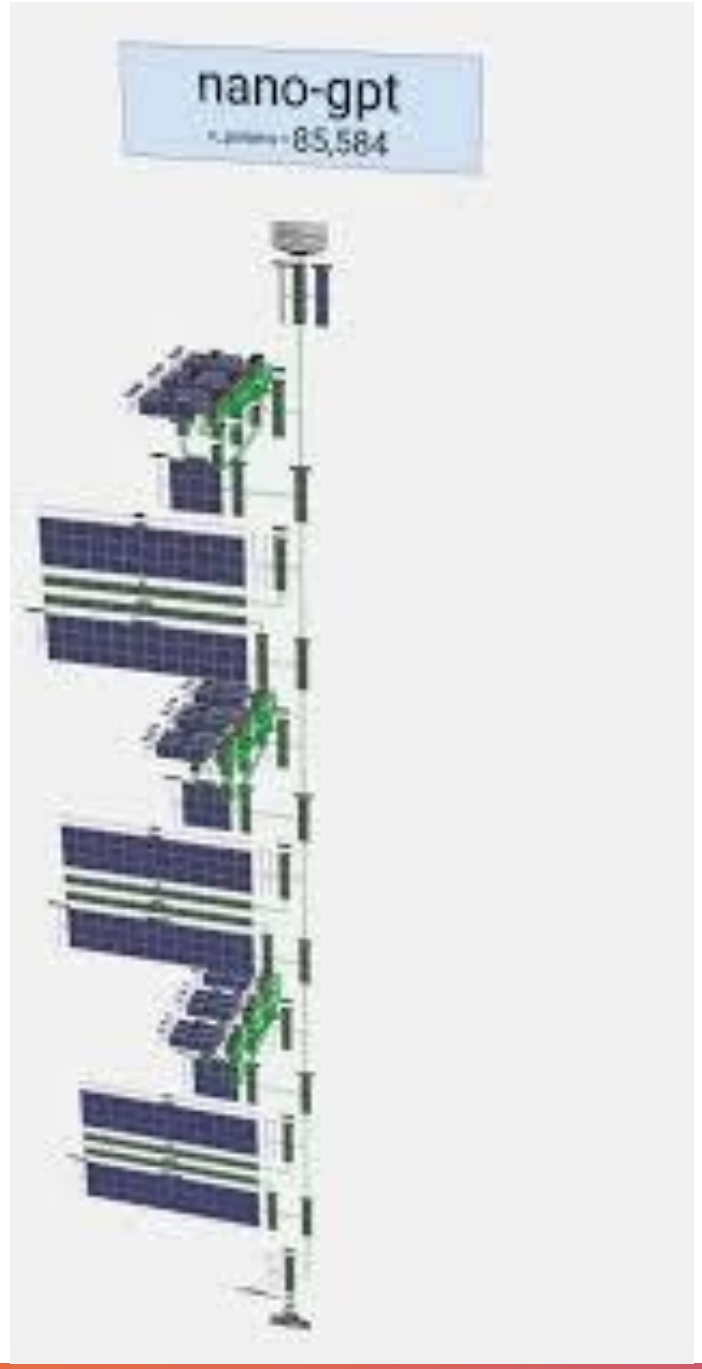




Se separó en codificadores y decodificadores de información



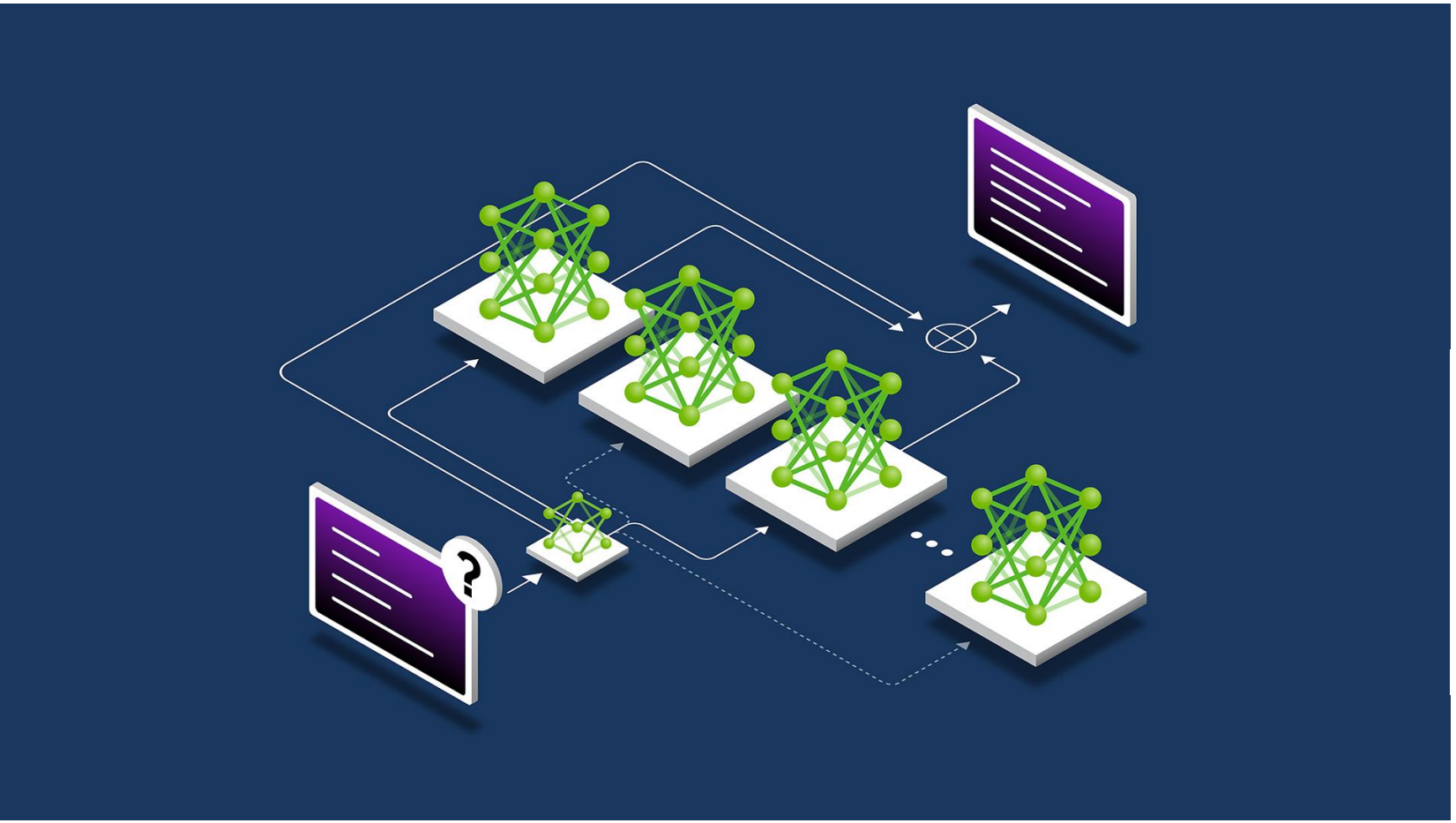
*Illustrative example, exact model architecture may vary slightly



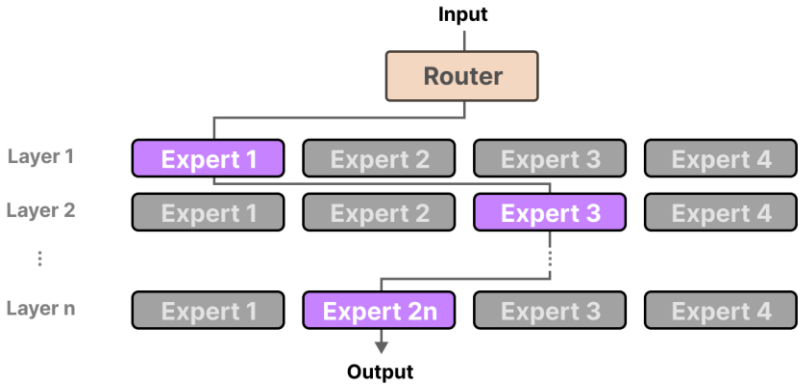
Se juntaron los bloques y boom! Nació un pequeño GPT

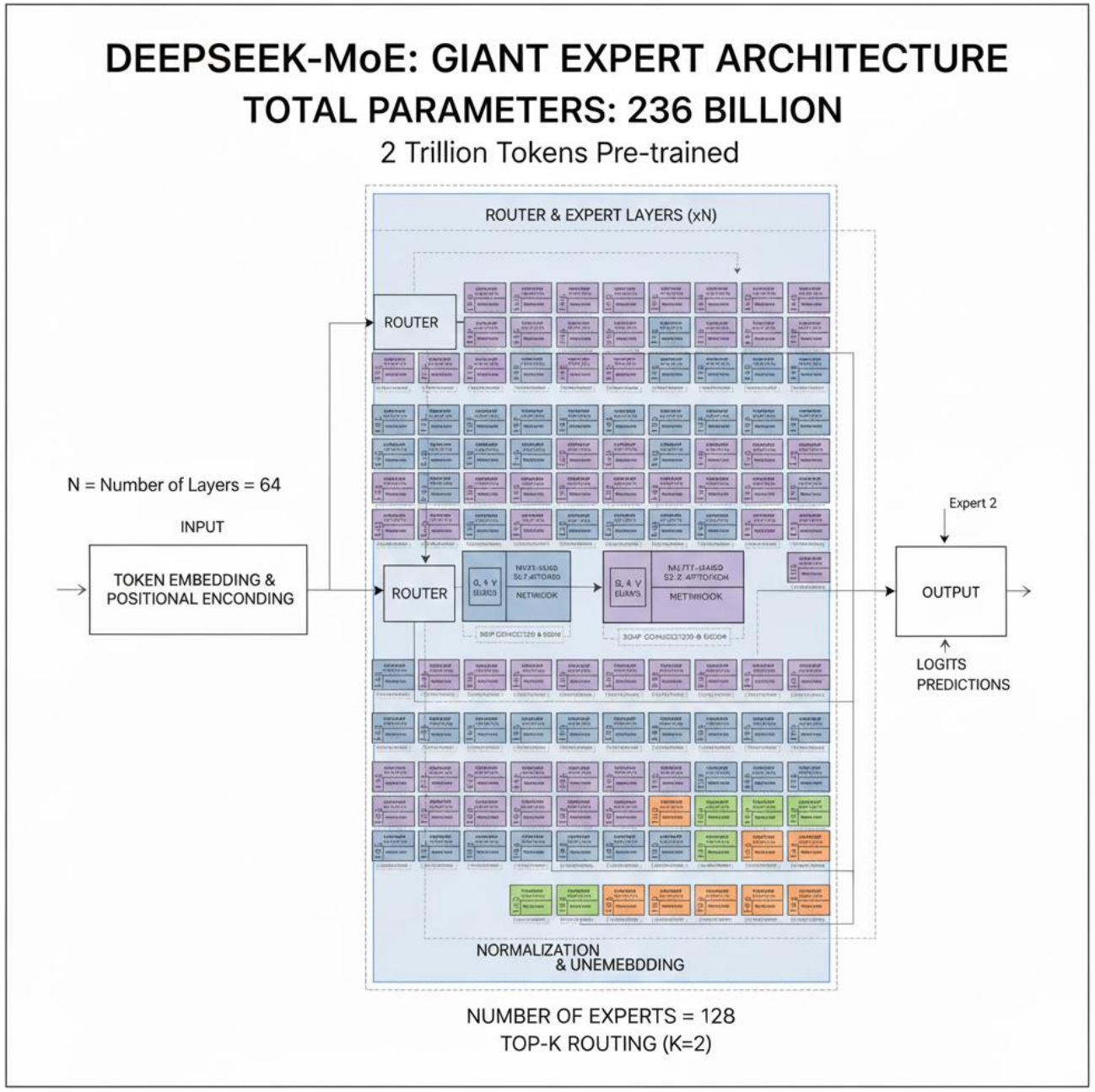
- NanoGPT tiene 3 bloques de 3 cabezales cada uno
- Son 85mil neuronas!

Y más y más grandes... 96 bloques de 96 cabezales cada uno... y nació GPT3

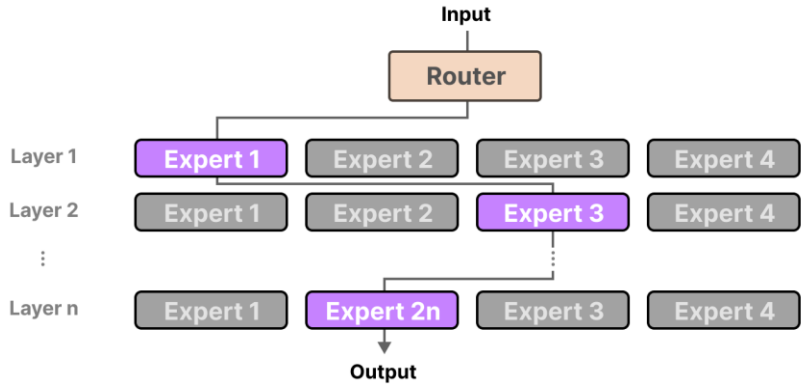


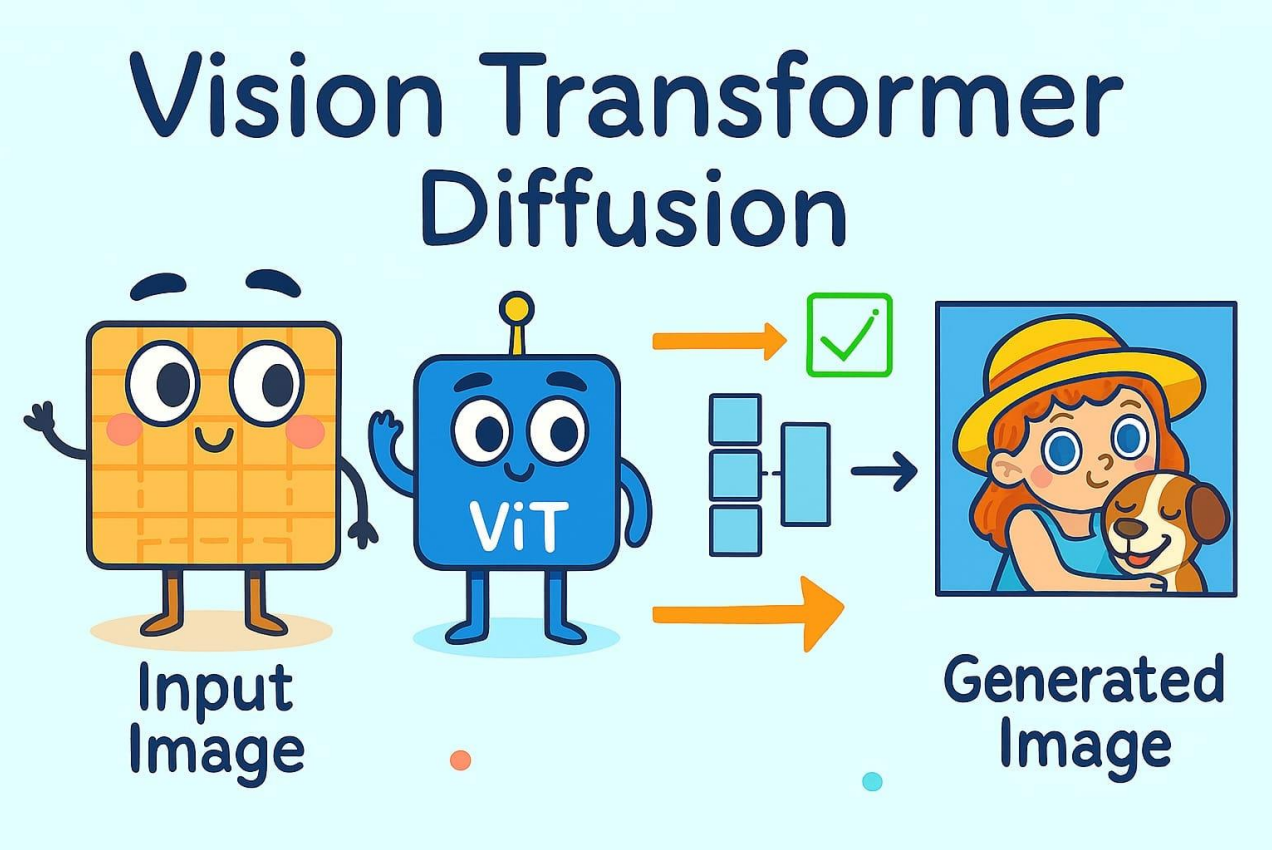
Ahora tienen cientos y miles de bloques con sub redes



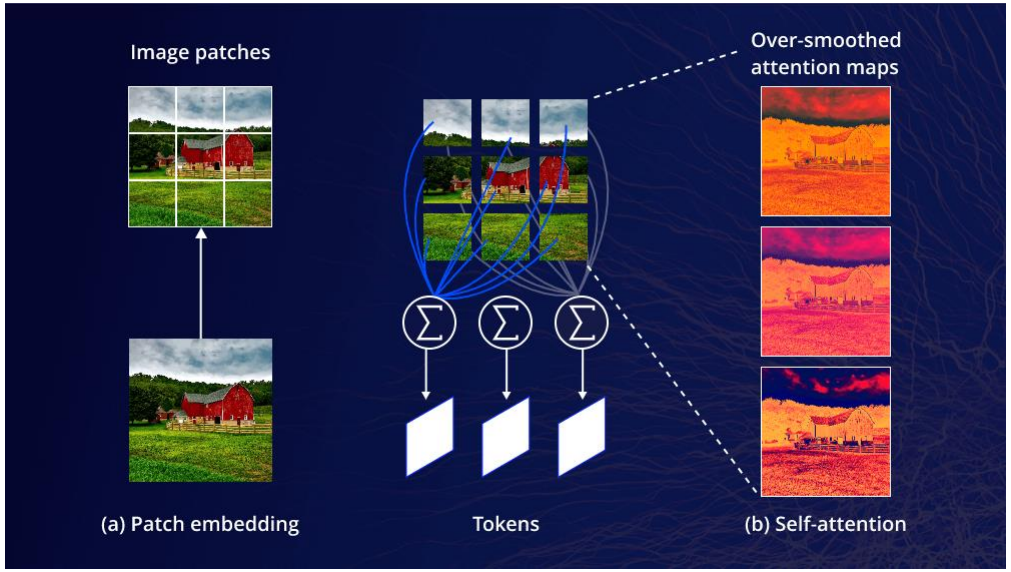


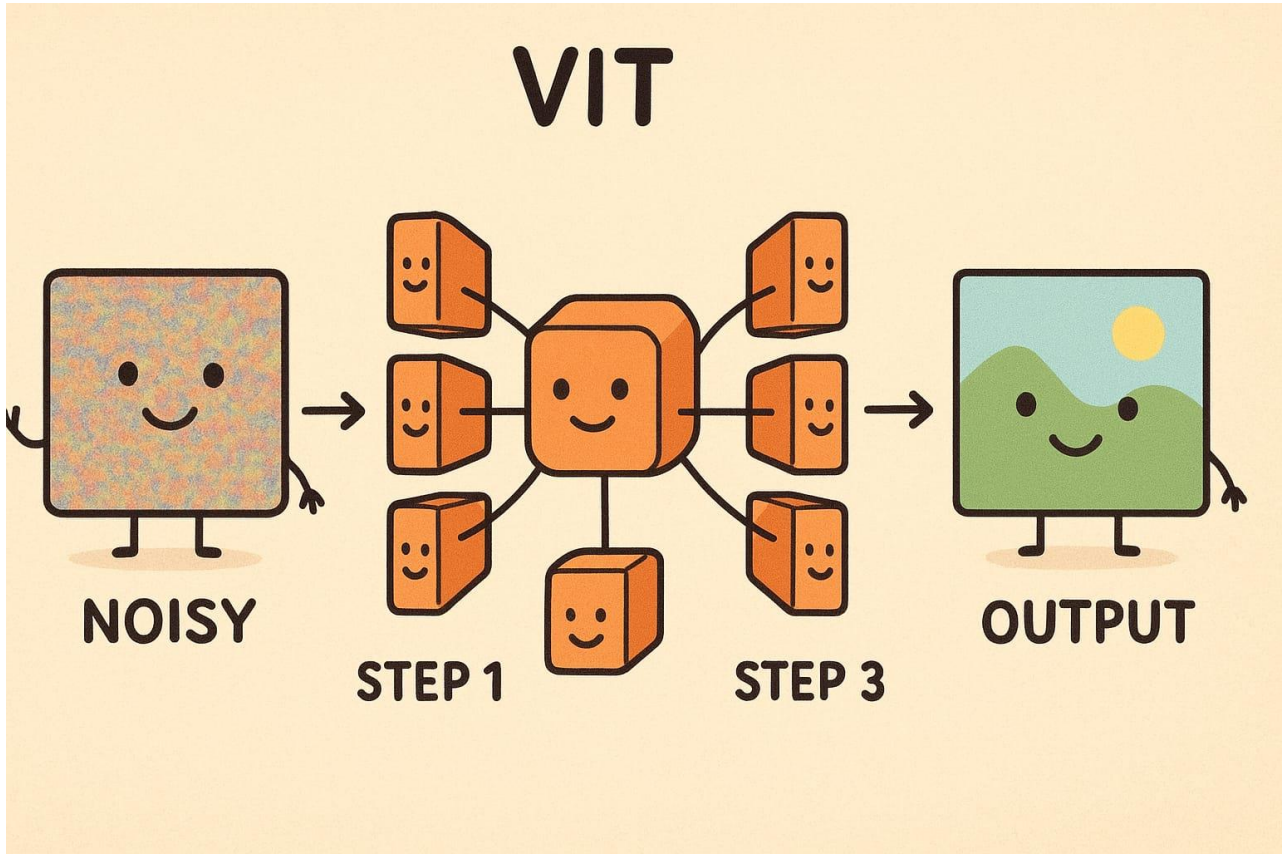
Cientos de expertos, cada uno con decenas de bloques con decenas de cabezales cada bloque



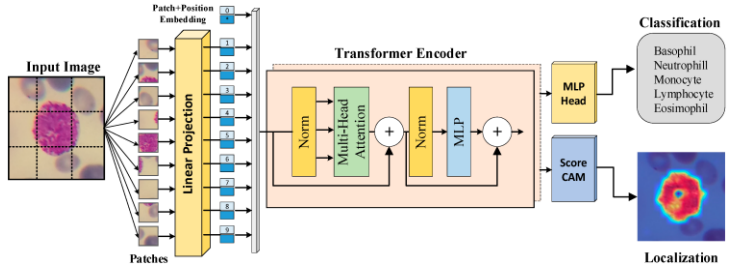


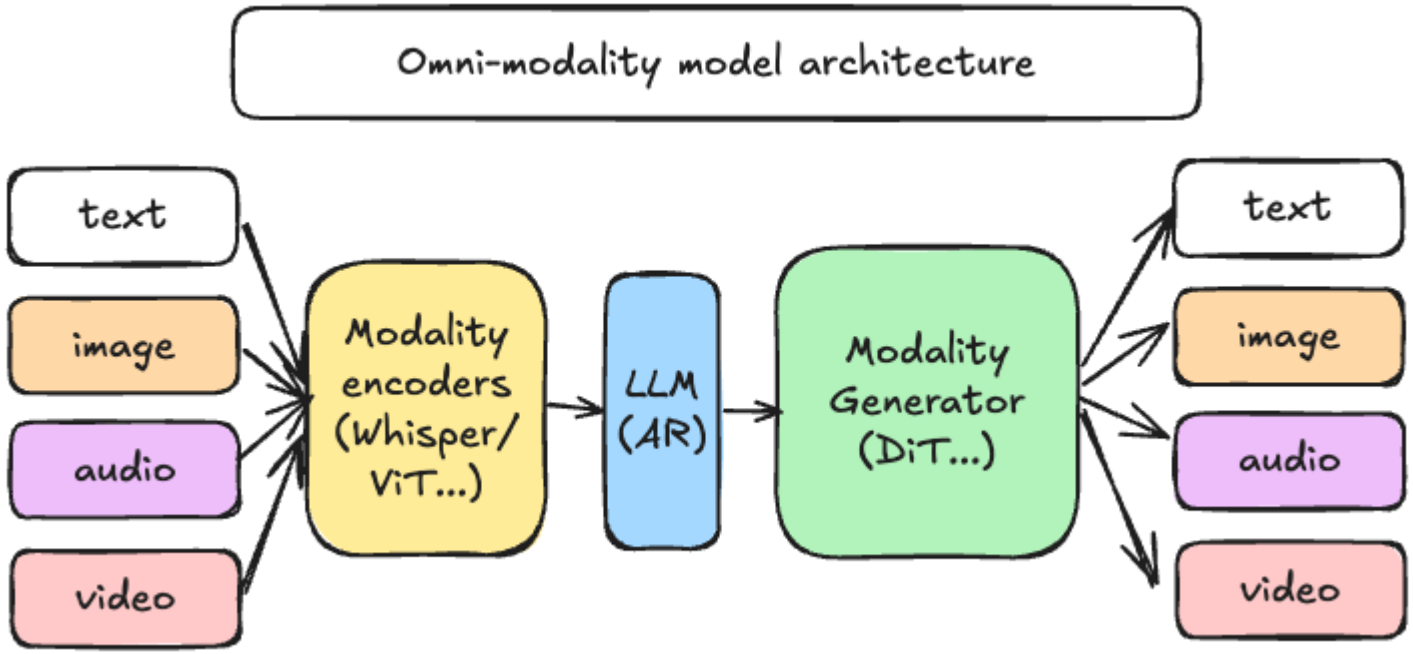
Y de pronto no sólo era texto, también imágenes



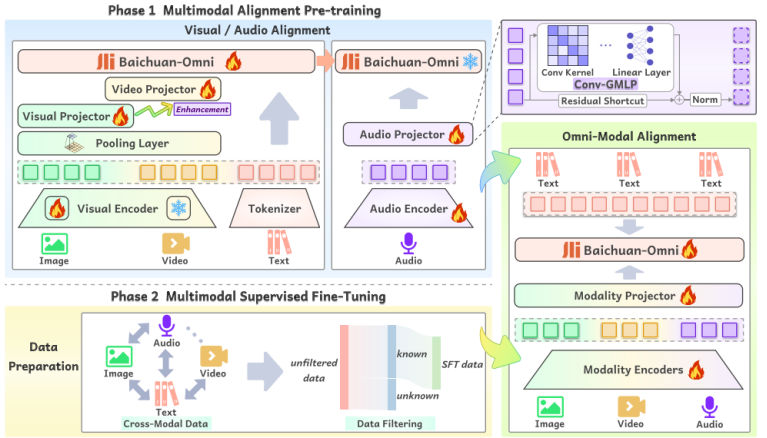


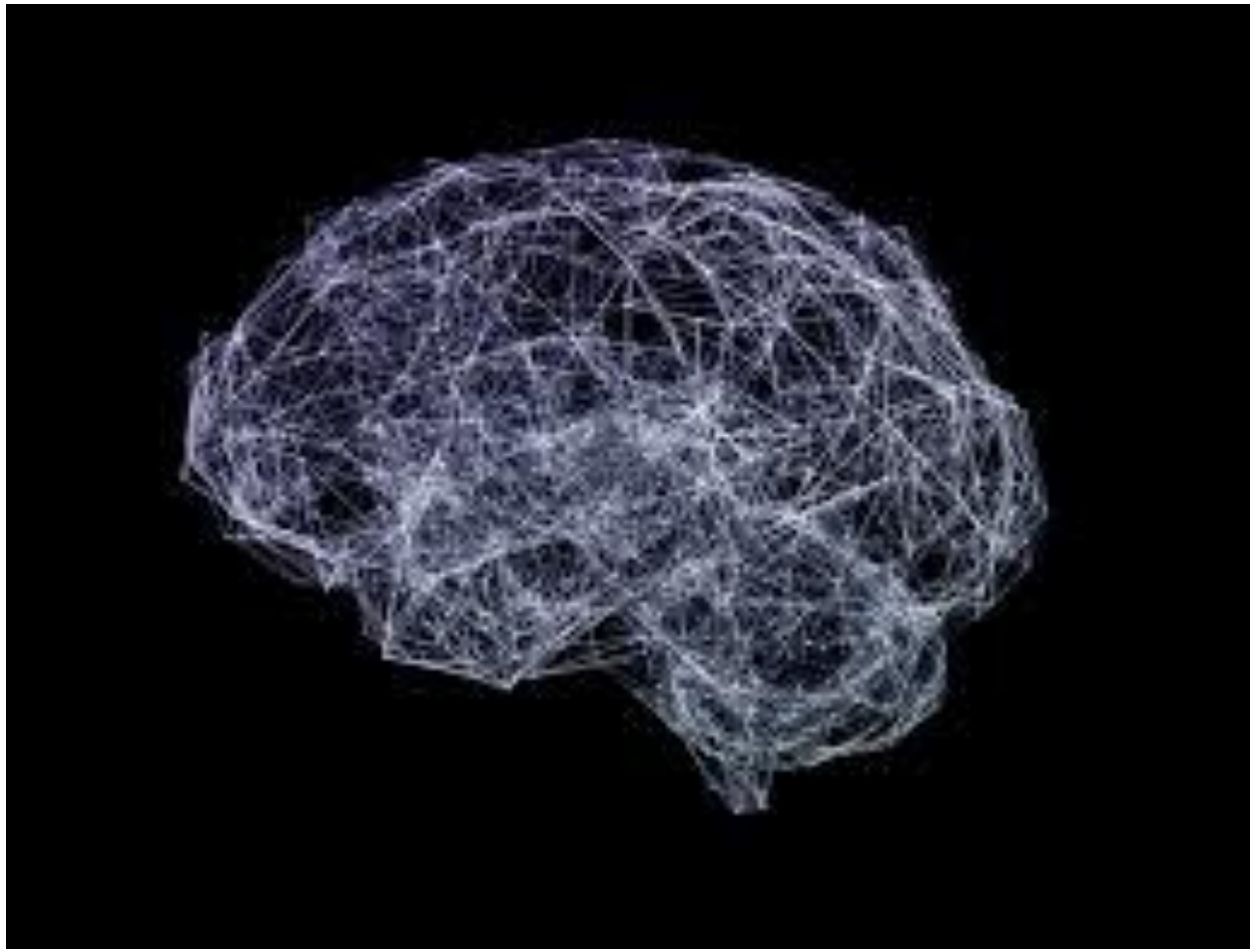
Y de pronto fuimos capaces de generar imágenes más y más perfectas





Finalmente, tenemos redes neuronales capaces de entender y generar de todo





El cerebro tiene entre 100T a 1000T de conexiones, conectadas en paralelo.
Los mejores modelos llegan a 1T o 2T, quizás a 5T los más caros, con conexiones secuenciales en cadena.
Aún nos falta mucho!