

Transformer

QUÉ HACE ESTE MINI-TRANSFORMER

★ Paso 1: Convertir tokens a vectores

Se crean representaciones simples:

- "yo" → [1, 0]
- "soy" → [0, 1]

★ Paso 2: Agregar positional encoding

Para distinguir las posiciones:

css

 Copy code

```
pos 0 → [0.0, 0.1]
pos 1 → [0.1, 0.0]
```

★ Paso 3: Self-Attention

Se calculan:

- **Q = query**
- **K = key**
- **V = value**

Después:

ini

 Copy code

```
scores = Q·KT
attn = softmax(scores)
Z = attn·V
```

Esto determina cuánto se atienden entre si "yo" y "soy".

★ Paso 4: Feed-Forward

Finalmente una mini-red lineal para cada token:

ini

 Copy code

```
ff = z·w1 + b1
```

Por qué este es el Transformer más simple posible

- No usa multi-head.
- No usa normalización.
- No usa entrenamiento.
- Todo son matrices 2×2 fáciles de validar a mano.
- Los números son tan pequeños que puedes pasar todo a Excel sin complicarte.