Positional Encoding

- Encode token position directly in the word vector:

POSITIONAL ENCODING	0 0 1 1	0.84 0.0001 0.54 1	0.91 0.0002 -0.42 1
	+	+	+
EMBEDDINGS	x_1	x_2	x_3
INPUT	Je	suis	étudiant
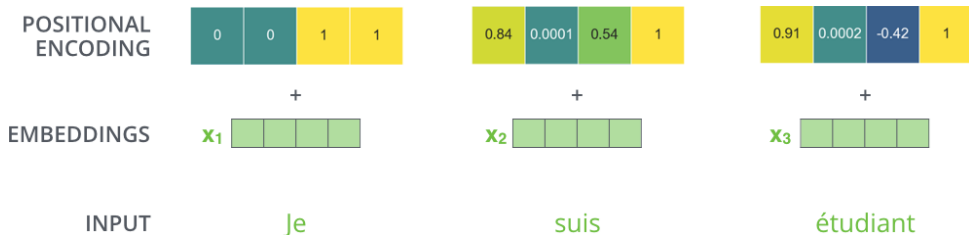
- Positional embedding can be random, or "frequency-like":