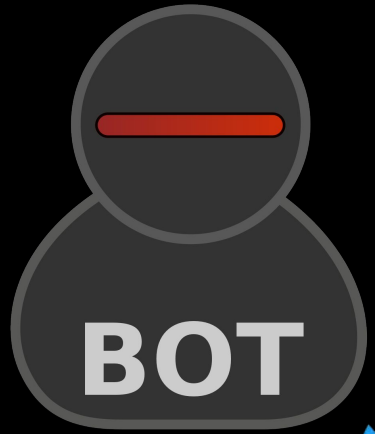
ML for
Natural Language Processing

Motivación


BOT

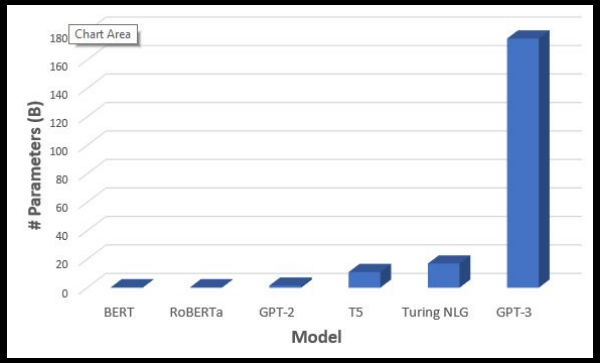
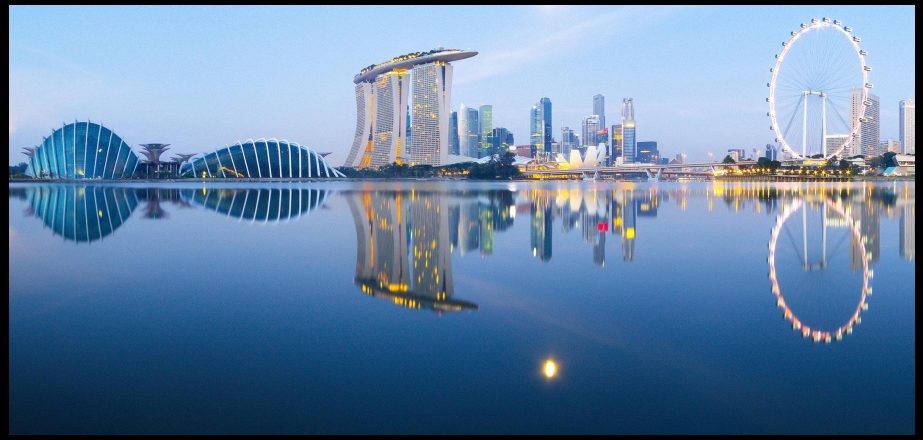
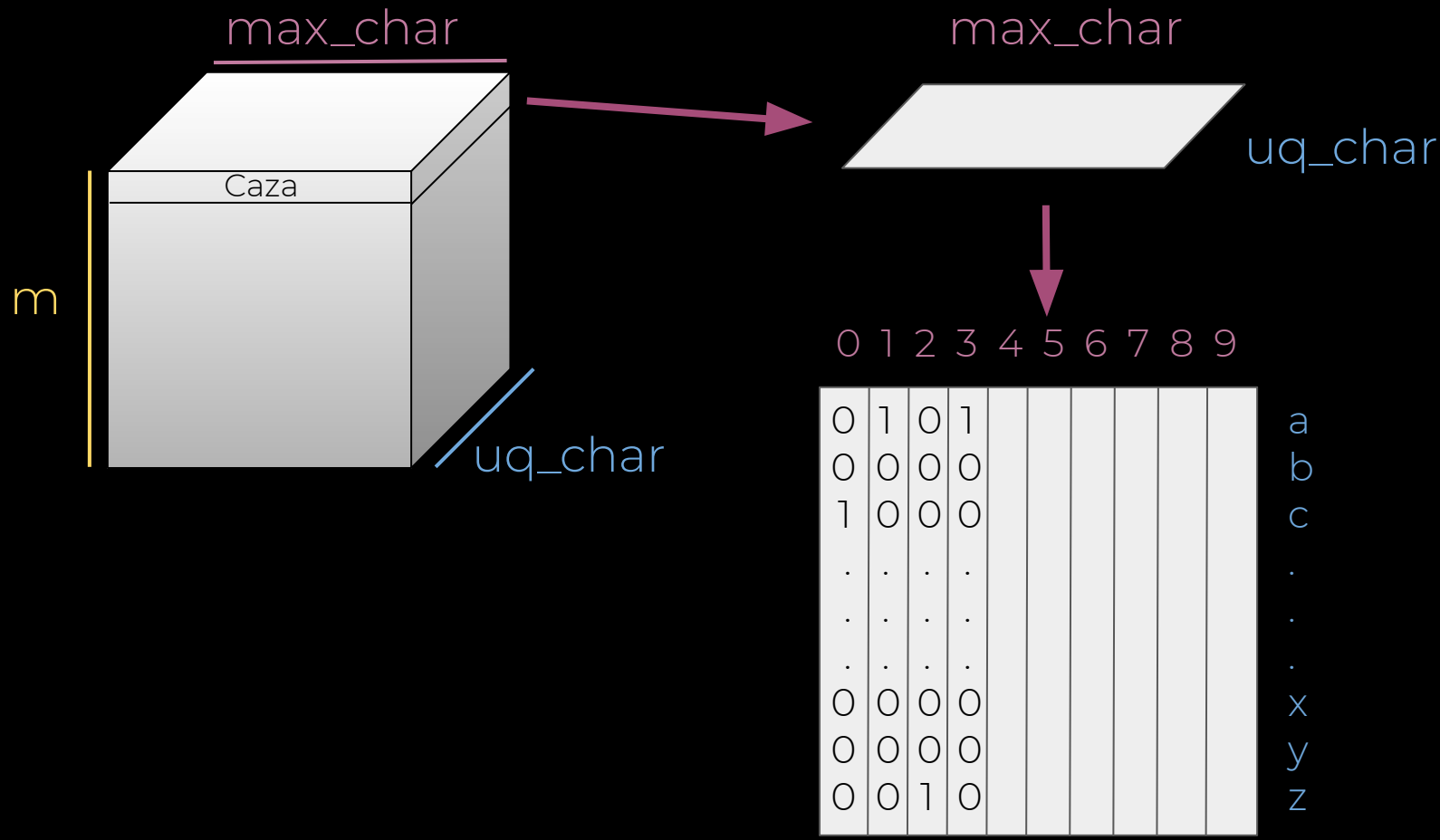


Chart Area
Parameters (B)
180
160
140
120
100
80
60
40
20
0
BERT
RoBERTa
GPT-2
T5
Turing NLG
GPT-3
Model

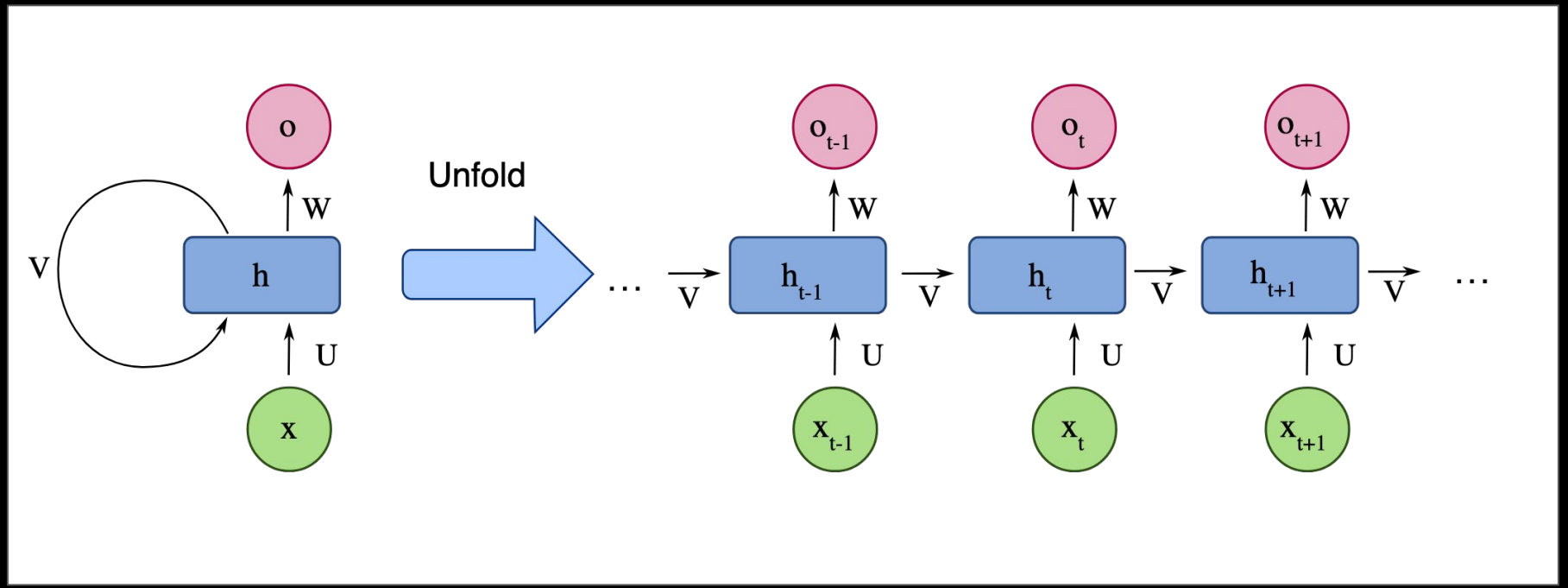
Character level language modelling



Let’s get to the point now.


max_char
Caza
m
uq_char
max_char
uq_char
0 1 2 3 4 5 6 7 8 9
0 1 0 1
0 0 0 0
1 0 0 0
. . . .
. . . .
. . . .
0 0 0 0
0 0 0 0
0 0 1 0
a
b
c
.
.
.
x
y
z

¿Qué es una RNN?



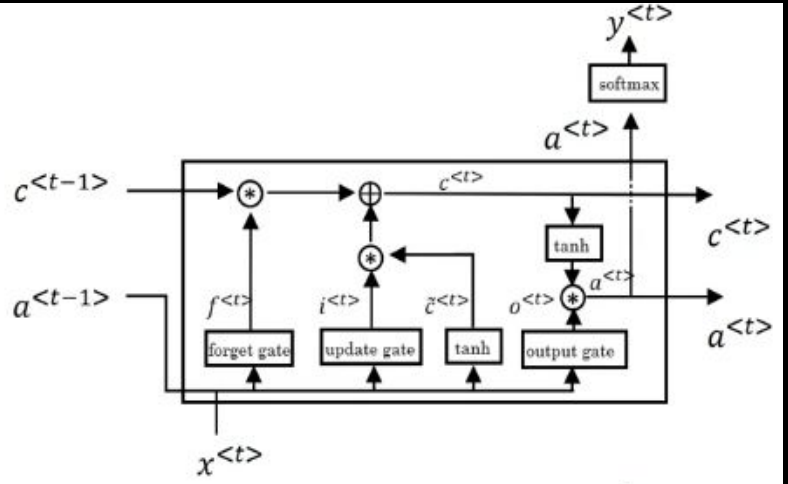
¿Qué hace una red neuronal recurrente?

P(carácter siguiente **|** caracteres previos**)**

vector de probabilidades = **(**0, 0, 0, 1, 0**)**

¿Qué es un LSTM?

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c)$$
$$\Gamma_u = \sigma(W_u[\, a^{<t-1>}, x^{<t>}] + b_u)$$
$$\Gamma_f = \sigma(W_f[\, a^{<t-1>}, x^{<t>}] + b_f)$$
$$\Gamma_o = \sigma(W_o[\, a^{<t-1>}, x^{<t>}] + b_o)$$
$$c^{<t>} = \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>}$$
$$a^{<t>} = \Gamma_o * \tanh c^{<t>}$$



Gracias.