

# Text as a sequence of characters

One-hot  
encoded  
characters,  
length  $\sim 70$

	<b>_</b>	<b>c</b>	<b>a</b>	<b>t</b>	<b>_</b>	<b>r</b>	<b>u</b>	<b>n</b>	<b>s</b>	<b>_</b>
	0	0	1	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0
	0	1	0	...	0	...	...	...	...	0
	...	...	...	1	...	1	1	1	1	...
	0	0	0	...	0	...	...	...	...	0

Let's start with character  $n$ -grams!