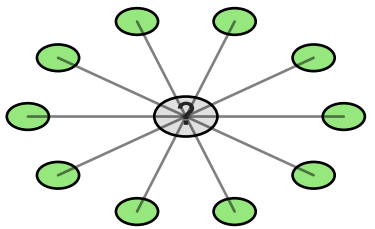
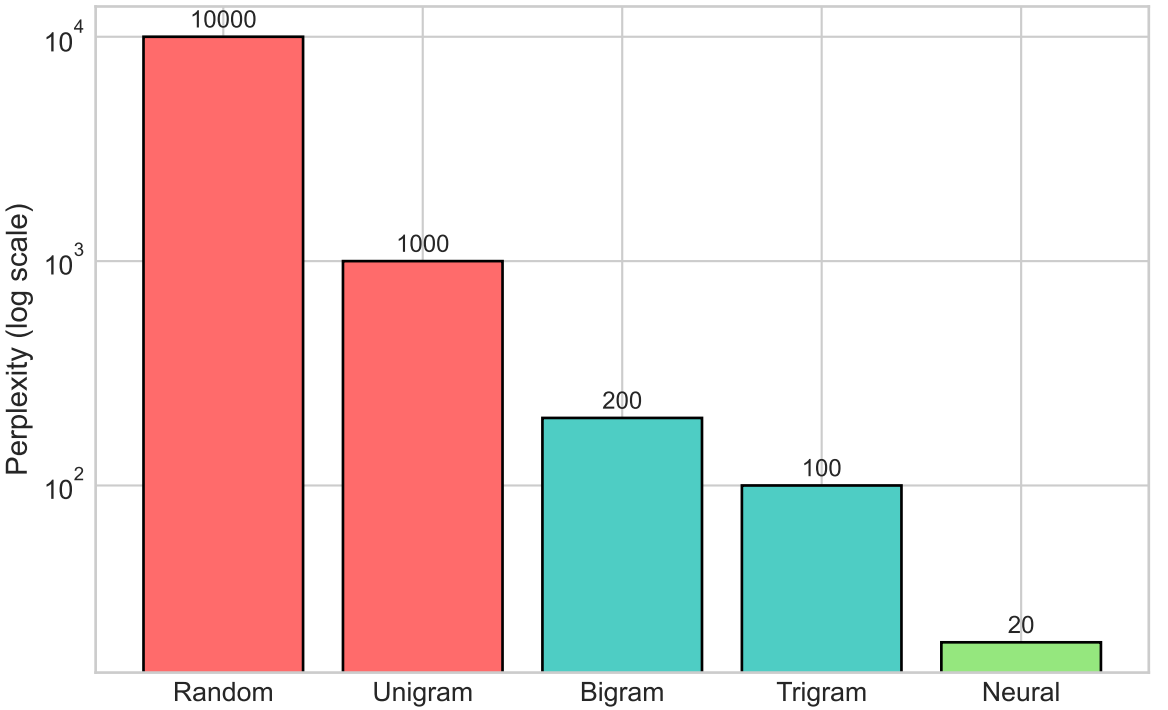
What is Perplexity?



Perplexity = 10

Model is choosing among
10 equally likely options

Model Comparison



The Math

$$\text{Perplexity} = 2^{H}$$

where $H = -1/N \sum \log_2 P(w_i|\text{context})$

Lower perplexity = Better model

Intuitive Examples

PP = 2	→	Like flipping a coin
PP = 6	→	Like rolling a die
PP = 26	→	Like choosing a letter
PP = 1000	→	Like choosing from dictionary