Trigram Model

$$p(w_1, w_2, w_3, \ldots w_n) \approx p(w_1)\, p(w_2 \mid w_1)\, p(w_3 \mid w_2, w_1) \ldots p(w_n \mid w_{n-1}, w_{n-2})$$

$$\text{Perplexity}(C) = N \sqrt{\frac{1}{\prod_{i=1}^{m} P(s_i)}}$$

$$= 2^{-\frac{1}{N} \sum_{i=1}^{m} \log_2 (P(s_i))}$$

$N$ = total length of test sentences

Sentence $s$ with $n$ words, $\sqrt[n]{\frac{1}{P(w_1^n)}}$.

$$\text{Perplexity}(C) = N \sqrt{\frac{1}{P(s_1, s_2 \ldots s_k)}}$$ Perplexity of corpus

$$\text{Perplexity}(C) = 2^{-\frac{1}{N} \sum_{i=1}^{k} \log_2 (p(w_1, w_2, w_3 \ldots w_n))}$$