

Robust Multi-bit Natural Language Watermarking through Invariant Features

<https://arxiv.org/pdf/2305.01904.pdf>

- 这篇文章思想非常简单，就是在KGW的基础上先算了一个熵，小于某个阈值就不加水印了，某种程度的提升文本质量

- 熵的计算：

$$H^{(t)} = - \sum_{i=1}^{|\mathcal{V}|} p_i^{(t)} \log p_i^{(t)}.$$