

Word Prediction: Formal Definition

Binary Random Variable : $X_w = \begin{cases} 1 & \text{w is present} \\ 0 & \text{w is absent} \end{cases}$
 $X_w \in \{0, 1\}$

$$p(X_w = 1) + p(X_w = 0) = 1$$

The more random X_w is, the more difficult the prediction would be.

How does one quantitatively measure the “randomness” of a random variable like X_w ?