

Enryk Elizondo González

① Técnica de Absolute Discounting

- Un valor constante "m" se sustrae de cada cuenta. El efecto de esto es que los eventos con menores cuentas son descontados relativamente más que aquellos con mayores cuentas.

El coeficiente de descuento d_a se calcula de la siguiente manera:

$$d_a = \frac{q - m}{q}$$

② Técnica de smoothing Kneser - Ney para n-gramas

- Se usa para calcular la probabilidad distribuida de n-gramas en un documento basado en sus historias.

$$P_{KN}(w_i | w_{i-n+1}^{i-1}) = \frac{\max(c_{KN}(w_{i-n+1}^{i-1}) - d, 0) + \lambda(w_{i-n+1}^{i-1})}{c_{KN}(w_{i-n+1}^{i-1})} * P_{KN}(w_i | w_{i-n+2}^{i-1})$$