

Absolute Discounting

Trata de redistribuir parte de la probabilidad de los N -gramas frecuentes a los N -gramas no observados.

- Expresión matemática: $P_{AD}(w_i | w_{i-n+1}^{i-1}) = \frac{\text{Count}(w_{i-n+1}^i)}{\sum \text{Count}(w_{i-n+1}^{i-1})}$

- Ejemplo: Si un trígrama como "El color rojo" aparece 10 veces y aplicamos un descuento de 0.75, la frecuencia sería 9.25 y el resto de la probabilidad se reasigna a trigramas observados.

Kneser - Ney smoothing

Además de reasignar probabilidades no observados, mejora la estimación para los N -gramas raros, especialmente aquellos que consisten en palabras comunes pero en combinaciones inusuales.

- Expresión matemática

$$P_{KN}(w_i | w_{i-n+1}^{i-1}) = \frac{\max(\text{Count}(w_{i-n+1}^i) - D, 0)}{\sum \text{Count}(w_{i-n+1}^{i-1})} + \lambda(w_{i-n+1}^{i-1})$$

- Ejemplo: Si se tiene un bigrama "de los" en lugar de solo contar, con Kneser-Ney también tiene en cuenta cuantas veces "los" sigue a diferentes palabras.