

Tecnológico de Monterrey  
Inteligencia artificial avanzada para la ciencia de datos II  
Gpo 101  
Investigación II de N-gramas  
Saúl Francisco Vázquez del Río A01198261

Tarea 4 Investigación II de N-grams

• Investigue a mano en qué consisten las técnicas de N-grams: absolute discounting y smoothing kneser-Ney qué problemática resuelven, algún breve ejemplo y la expresión matemática involucrada

1. Absolute discounting es un enfoque sencillo en el que se resta un discount fijo  $D$  del recuento de cada  $n$ -gram observado. La masa de la probabilidad restante luego se distribuye para tener cuenta los  $n$ -grams posibles

$$P(w_i | w_{i-n+1}^{i-1}) = \frac{c(w_i | w_{i-n+1}^{i-1}) - D}{c(w_i | w_{i-n+1}^{i-1})} + \lambda \cdot P(w_i | w_{i-n+2}^{i-1})$$

$$P(\text{sat} | \text{the}) = 0 + 0.575 \times 0.125 = 0.071875$$

The cat sat

$$\lambda = \frac{0.5 \cdot 6}{8} = \frac{3}{8} = 0.375$$

2. kneser-Ney es una de las técnicas de descuento más sofisticadas y utilizadas. Existe el absolute discount no sólo a las probabilidades de  $n$ -grams observados sino también la distribución de retroceso.

$$P_{KN}(w_i | w_{i-2}^{i-1}) = \max\left(\frac{c(w_{i-2}, w_{i-1}, w_i) - D}{c(w_{i-2}, w_{i-1})}, 0\right) + \lambda (w_{i-2}, w_{i-1})$$

The cat sat

$$P_{KN}(w_i | w_{i-1})$$

$$(cat | the) = \frac{\max(2 - 0.75, 0)}{8} + 0.5625 \times 0.0909 \times \frac{1.25}{8} + 0.5625 \times 0.0909 \times 0.15625 = 0.0502 \approx 0.2071$$