

Simplest VSM= Bit-Vector + Dot-Product + BOW

$$\begin{aligned} \mathbf{q} &= (x_1, \dots, x_N) & x_i, y_i &\in \{0, 1\} \\ \mathbf{d} &= (y_1, \dots, y_N) & 1: \text{word } W_i \text{ is present} \\ & & 0: \text{word } W_i \text{ is absent} \end{aligned}$$

$$\text{Sim}(\mathbf{q}, \mathbf{d}) = \mathbf{q} \cdot \mathbf{d} = x_1 y_1 + \dots + x_N y_N = \sum_{i=1}^N x_i y_i$$

What does this ranking function intuitively capture?
Is this a good ranking function?