

Adapting BM25 Retrieval Model for Paradigmatic Relation Mining

$$\mathbf{d1}=(x_1, \dots x_N) \quad \text{BM25}(w_i, \mathbf{d1}) = \frac{(k+1)c(w_i, \mathbf{d1})}{c(w_i, \mathbf{d1}) + k(1-b + b*|\mathbf{d1}|/avdl)}$$

$$x_i = \frac{\text{BM25}(w_i, \mathbf{d1})}{\sum_{j=1}^N \text{BM25}(w_j, \mathbf{d1})}$$

$$b \in [0, 1]$$

$$k \in [0, +\infty)$$

$$\mathbf{d2}=(y_1, \dots y_N) \quad y_i \text{ is defined similarly}$$

$$\text{Sim}(\mathbf{d1}, \mathbf{d2}) = \sum_{i=1}^N \text{IDF}(w_i) x_i y_i$$