

If I only had a vector: The Analogy Problem

To solve analogy problems of the form “ w_a is to w_b as w_c is to what?”, we can simply compute a query vector

$$\mathbf{q} = \mathbf{w}_b - \mathbf{w}_a + \mathbf{w}_c$$

and find the most similar word vector $\mathbf{v} \in \mathbf{W}$ to \mathbf{q} . If we normalize \mathbf{q} to unit-length

$$\hat{\mathbf{q}} = \frac{\mathbf{q}}{\|\mathbf{q}\|}$$

and assume each vector in \mathbf{W} is also unit-length, this reduces to computing

$$\arg \max_{\mathbf{v} \in \mathbf{W}} \mathbf{v} \cdot \hat{\mathbf{q}}$$

and returning the associated word v .