Corpus

"...the stock market..."

Sliding window

Efficiency Gain:

k+1 samples

instead of |V|

(k ≈ 5-20)

market

P(+) → 1

positive sample

stock

(target word)

negative samples

piano

dog

blue

chair

P(+) → 0

Objective:

$$\log \sigma(\mathbf{v}'_{context} \cdot \mathbf{v}_{target}) + \sum_{i=1}^{k} \log \sigma(-\mathbf{v}'_{neg_i} \cdot \mathbf{v}_{target})$$

Negative Sampling: Binary Classification Instead of Full Softmax