Softmax in Neural Language Models: From Scores to Word Probabilities

Input Scores

Transformation

Output Probabilities

king
3.2

queen
2.8

man
1.5

woman
1.3

child
0.2

SOFTMAX FUNCTION

$$p_i = \frac{e^{s_i}}{\sum_j e^{s_j}}$$

king
0.487

queen
0.327

man
0.089

woman
0.073

child
0.024

Σ = 1.000