The Neo popularity rank for web-pages

Maxim Zakharov <dp.maxime@gmail.com>
29 January 2007*

This article describes the *Neo Popularity Rank* that uses in the *DataparkSearch Engine*[1] tool to assist web-pages ordering.

To construct the *Neo* popularity rank, we assumes that web-pages are neurons, and links between web-pages are connections between neurons. Thus, for given neural network we can use one of training methods, the back-error propagation method. For this model, the popularity rank of a web-page is the activity level of corresponding neuron.

For simplicity of calculation, every indexed document (U) is considered separately. With all documents referring it ($I_1, I_2, \ldots I_n$) and also all documents, to which it refers ($O_1, O_2, \ldots O_m$), it forms a one-layered neural network:
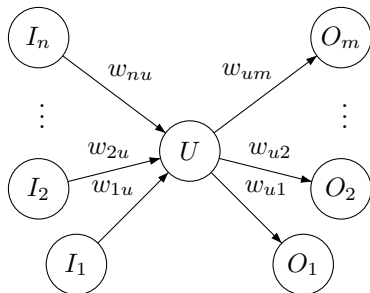
where $w_{1u}, w_{2u}, \ldots w_{nu}$ – are the weights for inbound links, and $w_{u1}, w_{u2}, \ldots w_{um}$ – are the weights for outbound links.

As an activation function we use well known sigmoid function

$$f(x) = \frac{1}{1 + e^{-x}}$$

The real activity level for neuron U is

$$o_u = f\left(\sum_{k=1}^{n} w_{ku} \cdot Neo(I_k)\right),$$

where $Neo(X)$ – is the value of popularity rank for neuron X.

As the expected activity level of neuron U we consider the value

$$d_u = f\left(\sum_{k=1}^{m} w_{uk} \cdot Neo(O_k)\right).$$

If a real and expected activity levels aren't equal, correction of weights of inbound links is made as follows:

$$\Delta w_{uk} = -c \cdot (d_u - o_u) \cdot o_u \cdot (1 - o_u) \cdot Neo(O_k), k = 1 \ldots n,$$

where c – is a constant factor of training speed. [1, Chapter 10]

In general, the number of iterations for a convergence $o_u \rightarrow d_u$ is not limited by any finite number, therefore after the fixed number of iterations set as parameter of algorithm, the value of *Neo*-rank for neuron U is set to

$$Neo(U) = \frac{d_u + o_u}{2}.$$

The steps of *Neo*-rank calculation for one neuron described above are consecutively applies to all indexed documents in a collection.

If some document has no inbound or outbound links, in this case the real or expected activity level for corresponding neuron is set equal to a some small value ε.

References

[1] George F. Luger, Artificial Intelligence: Structures and Strategies for Complex Problem Solving, Fourth Edition, Pearson Education Limited, 2002. ISBN 0-201-64866-0

---

*Last-Modified: 11 April 2007.
[1] See http://www.dataparksearch.org/