The Graph Neural Network Model

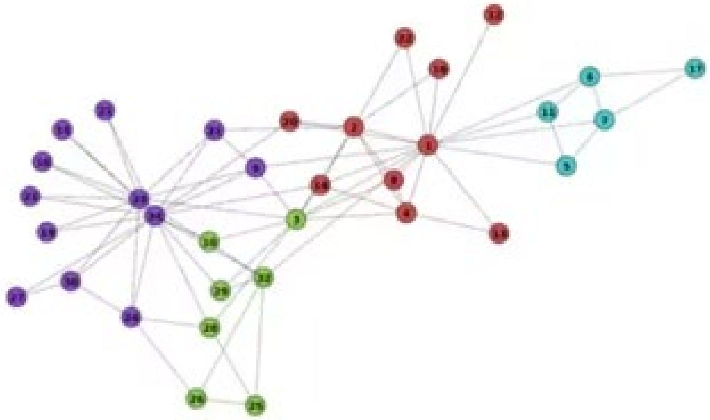
Franco Scarselli, Marco Gori, Fellow, IEEE, Ah Chung Tsoi, Markus Hagenbuchner, Member, IEEE, and Gabriele Monfardini

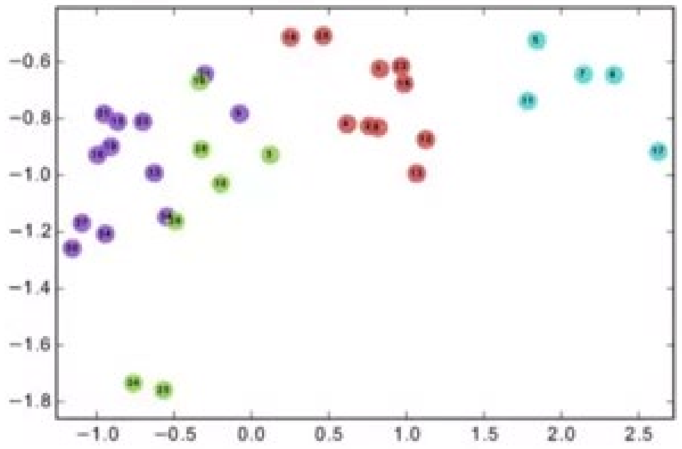
IEEE Trans

Contents

Introduction

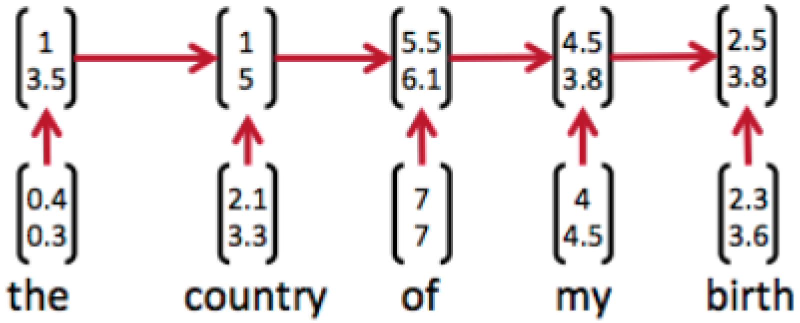
Traditional machine learning applications cope with graph structured data by using a preprocessing phase which maps the graph structured information to a simpler representation.

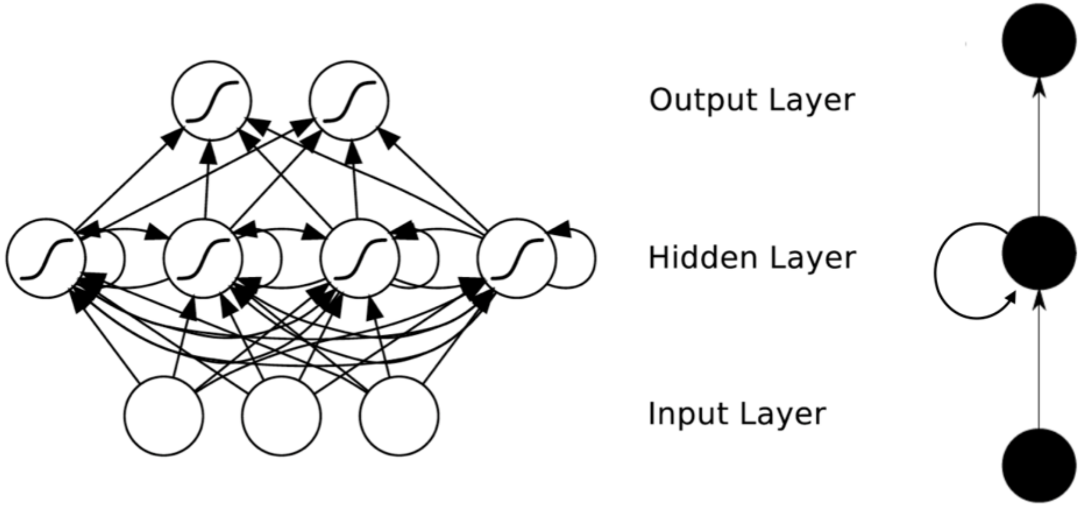




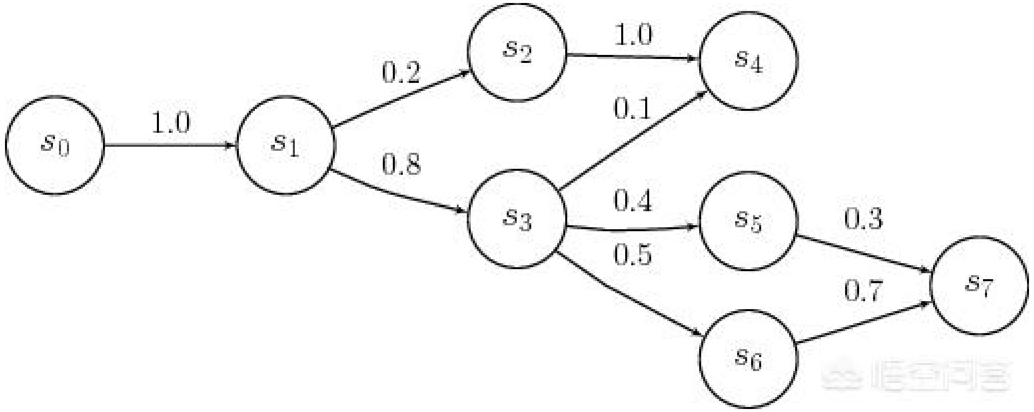
The idea is to encode the underlying graph structured data using the topological relationships among the nodes of the graph, in order to incorporate graph structured information in the data processing step. Recursive neural networks and Markov chains

Recursive neural networks





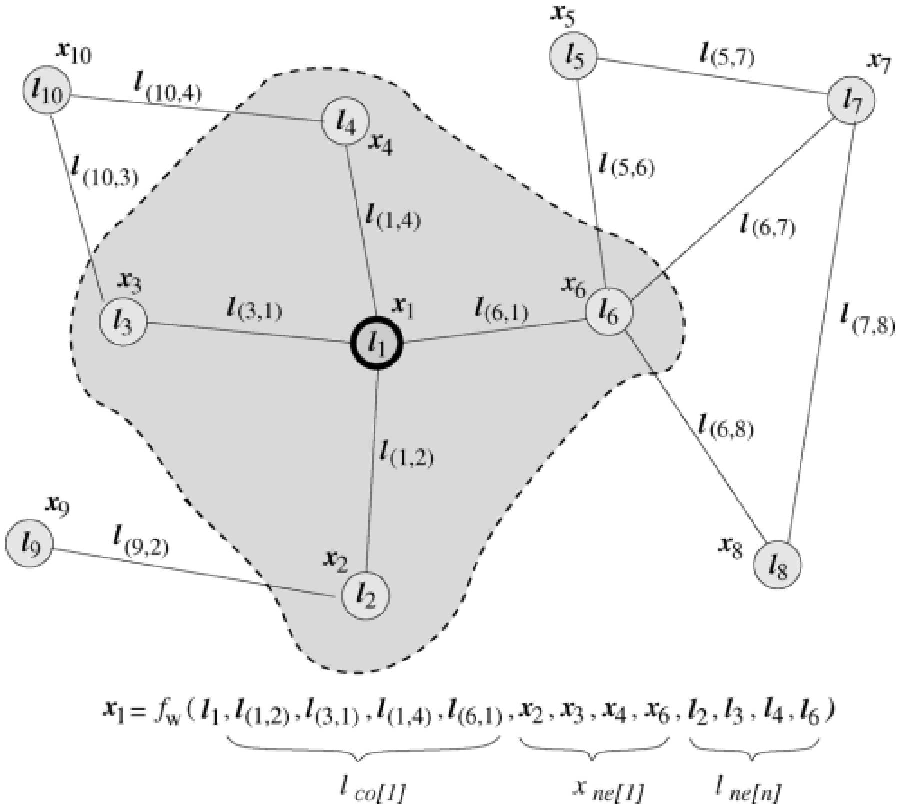
Markov chains



GNN model： Graph Neural Network Model

The intuitive idea underlining the proposed approach is that nodes in a graph represent objects or concepts, and edges represent their relationships. Each concept is naturally defined by its features and the related concepts. Thus, we can attach a state to each node that is based on the information contained in the neighborhood of . The state Xn contains a representation of the concept denoted by and can be used to produce an output , i.e., a decision about the concept

$$
\begin{aligned}
\boldsymbol{x}_n &= f_{\boldsymbol{w}}(\boldsymbol{l}_n, \boldsymbol{l}_{\mathrm{co}[n]}, \boldsymbol{x}_{\mathrm{ne}[n]}, \boldsymbol{l}_{\mathrm{ne}[n]}) \\
\boldsymbol{o}_n &= g_{\boldsymbol{w}}(\boldsymbol{x}_n, \boldsymbol{l}_n)
\end{aligned} \tag{1}
$$



Let X, O ,L,LN and be the vectors constructed by stacking all the states, all the outputs, all the labels, and all the node labels,respectively. Then, (1) can be rewritten in a compact form as

$$
\begin{aligned}
\boldsymbol{x} &= F_{\boldsymbol{w}}(\boldsymbol{x}, \boldsymbol{l}) \\
\boldsymbol{o} &= G_{\boldsymbol{w}}(\boldsymbol{x}, \boldsymbol{l}_N)
\end{aligned} \qquad (2)
$$

We are interested in the case when x,o are uniquely defined and (2) defines a map , which takes a graph as input and returns an output o for each node

The Banach fixed point theorem

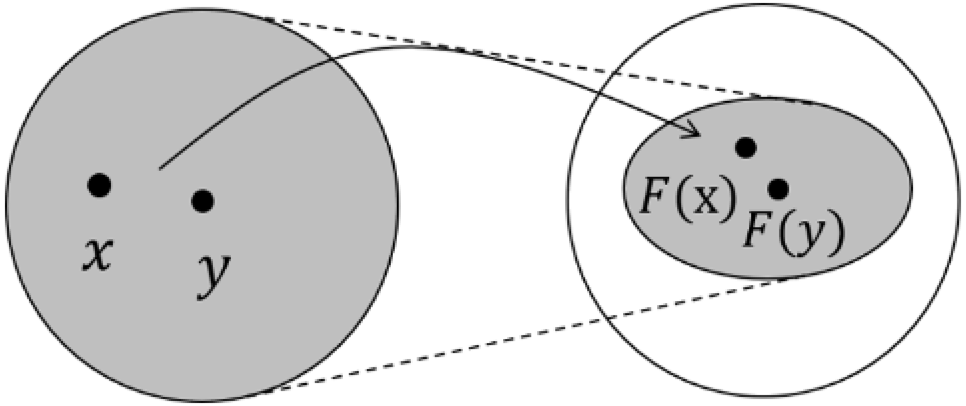
The Banach fixed point theorem

The moving point theorem means that no matter what X is, as long as Fw is a Compression map, X will converge to a fixed point after repeated iterations. We call it a fixed point.

$$\boldsymbol{x}(t+1) = F_{\boldsymbol{w}}(\boldsymbol{x}(t), \boldsymbol{l}) \qquad (4)$$

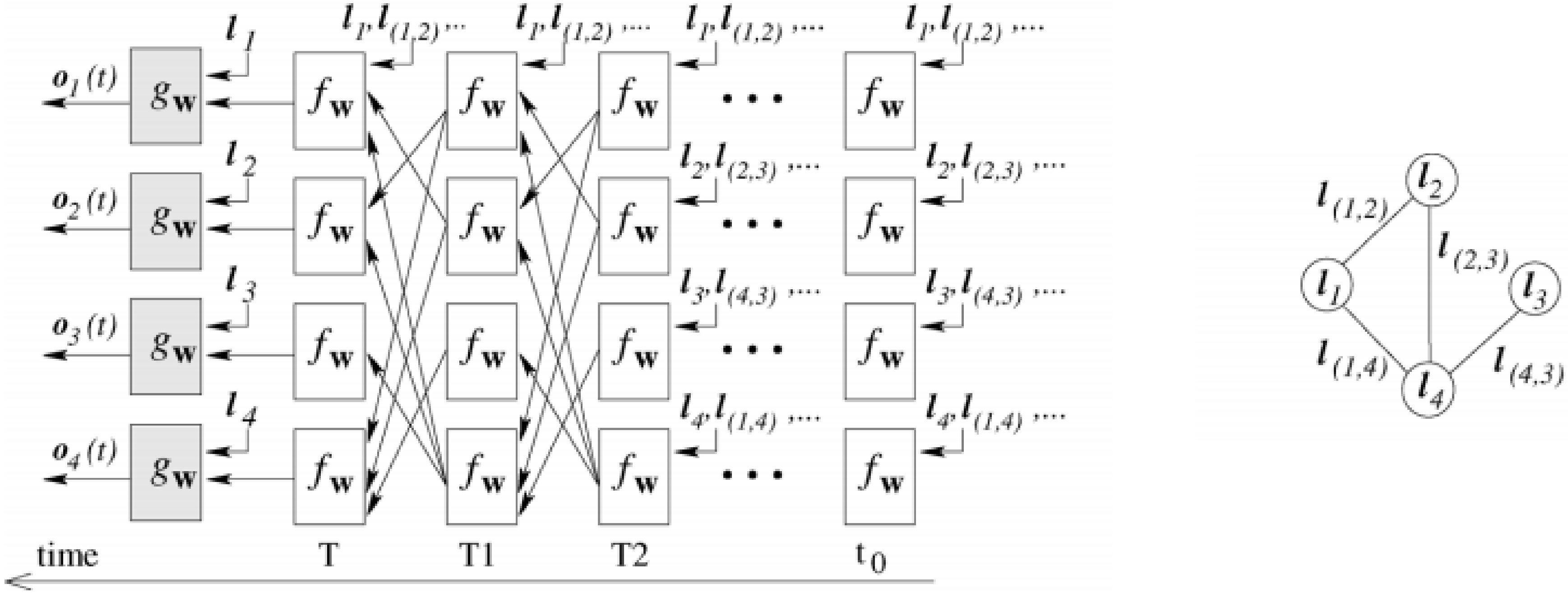
$$\boldsymbol{x}(t) = F_{\boldsymbol{w}}(\boldsymbol{x}(t), \boldsymbol{l})$$

Compression map



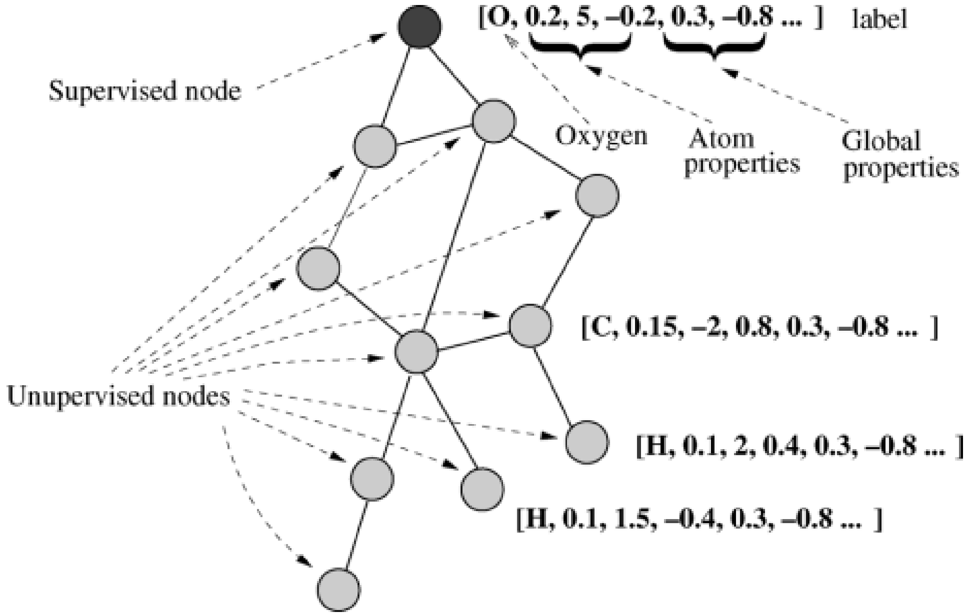
$$d(F(x), F(y)) \leq c\,d(x, y)$$
$$0 \leq c < 1$$

$$
\begin{aligned}
\boldsymbol{x}_n(t+1) &= f_{\boldsymbol{w}}(\boldsymbol{l}_n, \boldsymbol{l}_{\mathrm{co}[n]}, \boldsymbol{x}_{\mathrm{ne}[n]}(t), \boldsymbol{l}_{\mathrm{ne}[n]}) \\
\boldsymbol{o}_n(t) &= g_{\boldsymbol{w}}(\boldsymbol{x}_n(t), \boldsymbol{l}_n), \qquad n \in \boldsymbol{N}.
\end{aligned} \tag{5}
$$



Expriments

The Mutagenesis Problem



Method	Features	Reference	Accuracy
non–linear GNN	AB+C+PS		94.3
Neural Networks	C+PS	[13]	89.0%
P-Progol	AB+C	[13]	82.0%
P-Progol	AB+C+FG	[13]	88.0%
MFLOG	AB+C	[84]	95.7%
FOIL	AB	[85]	76%
boosted-FOIL	not available	[86]	88.3%
$1nn(d_m)$	AB	[87]	83
$1nn(d_m)$	AB+C	[87]	91%
RDBC	AB	[88]	84%
RDBC	AB+C	[88]	83%
RSD	AB+C+FG	[89]	92.6%
SINUS	AB+C+FG	[89]	84.5%
RELAGGS	AB+C+FG	[89]	88.0%
RS	AB	[90]	88.9%
RS	AB+FG	[90]	89.9%
RS	AB+C+PS+FG	[90]	95.8%
SVM_P	not available	[91]	91.5

TABLE V

ACCURACIES ACHIEVED ON THE REGRESSION-UNFRIENDLY PART OF THE MUTAGENESIS DATA SET. THE TABLE DISPLAYS THE METHOD, THE FEATURES USED TO MAKE THE PREDICTION, AND A POSSIBLE REFERENCE TO THE PAPER WHERE THE RESULT IS DESCRIBED

Method	**Knowledge**	**Reference**	**Accuracy**
non–linear GNN	AB+C+PS		96.0%
$1nn(d_m)$	AB	[87]	72%
$1nn(d_m)$	AB+C	[87]	72%
TILDE	AB	[92]	85%
TILDE	AB+C	[92]	79%
RDBC	AB	[88]	79%
RDBC	AB+C	[88]	79%

TABLE VI

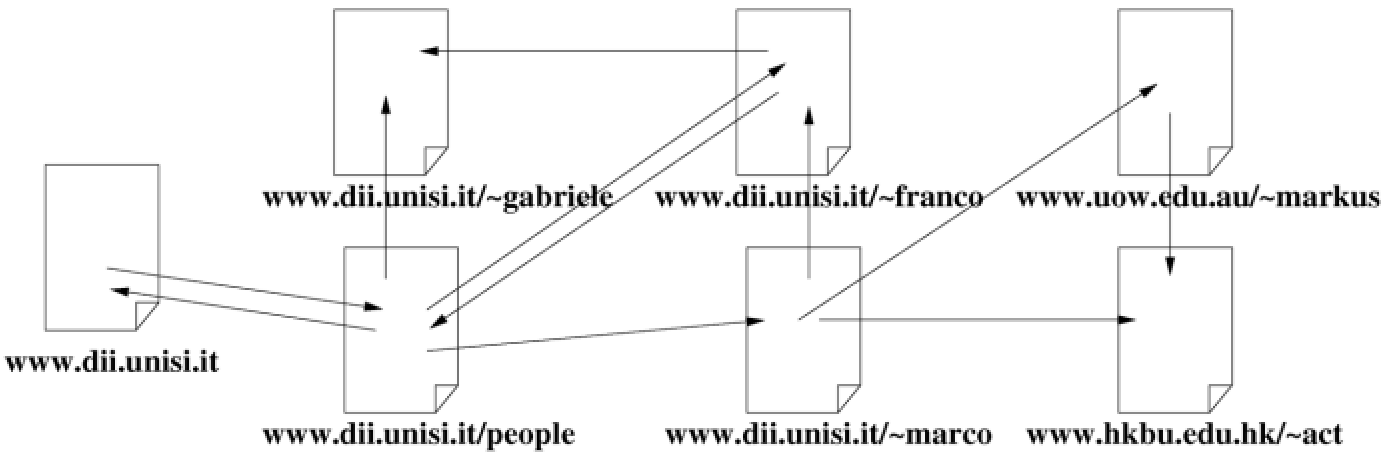
ACCURACIES ACHIEVED ON THE WHOLE MUTAGENESIS DATA SET. THE TABLE DISPLAYS THE METHOD, THE FEATURES USED TO MAKE THE PREDICTION, AND A POSSIBLE REFERENCE TO THE PAPER WHERE THE RESULT IS DESCRIBED

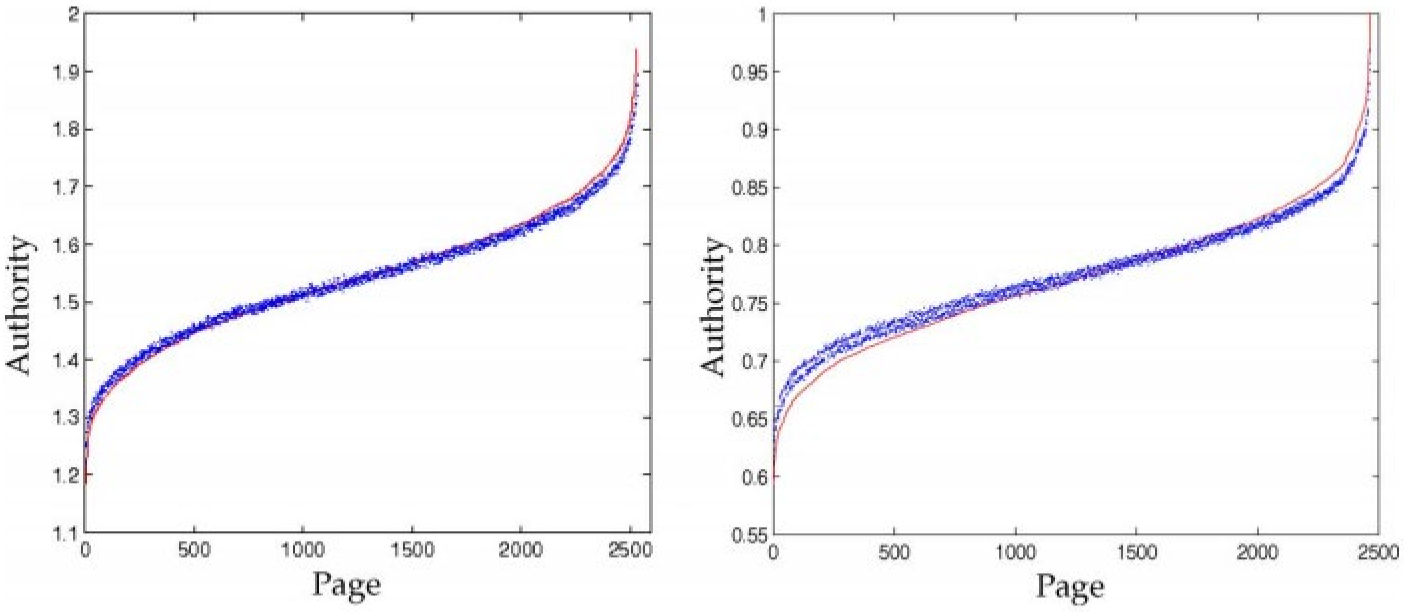
Method	**Knowledge**	**Reference**	**Accuracy**
non–linear GNN	AB+C+PS		90.5%
$1nn(d_m)$	AB	[87]	81%
$1nn(d_m)$	AB+C	[87]	88%
TILDE	AB	[92]	77%
TILDE	AB+C	[92]	82%
RDBC	AB	[88]	83%
RDBC	AB+C	[88]	82%

Web Page Ranking

According to PageRank, a page is considered authoritative if it is referred by many other pages and if the referring pages are authoritative themselves.

$$\tau(\boldsymbol{G}, n) = \begin{cases} 2p_n/\|\boldsymbol{p}\|_1, & \text{if } (a_n \,\mathrm{XOR}\, b_n) = 1 \\ p_n/\|\boldsymbol{p}\|_1, & \text{otherwise} \end{cases}$$





(a) (b)