Fast and Optimal Algorithms for Weighted Graph Prediction

**Context:** Analysis of Networked Data (hyperlinked webpages, social networks, co-author networks, biological networks, ...)

**Learning problem:** On-line Node Classification

**Basic Inductive Principle:** Strongly linked entities tend to belong to the same class

**On-line protocol:** Nodes are issued one by one in arbitrary order. At time $t$: learner predicts binary label $y_t$. Then $y_t$ is revealed

**Main issues:** (i) Accuracy performance guarantees (optimality), (ii) scalability and (iii) good practical performance

**Complexity Measure:** Expected number of cut-edges included in a uniformly generated random spanning tree

Nicolò Cesa-Bianchi †, Claudio Gentile ‡, Fabio Vitale †, Giovanni Zappella †

† Università degli Studi di Milano  ‡ Università degli Studi dell’Insubria
Weighted Tree Algorithm (WTA): Nearest Neighbor based on resistance distance metric on a spanning tree of the graph.

Accuracy Analysis: Optimal (up to log factors) on any weighted graphs.

Complexity
Time: Constant (amortized) per prediction
Space: Linear in graph size

Preliminary experiments: We compare WTA with fast global and local algorithms for on-line prediction on weighted graphs on two real world (biological) datasets.
WTA always outperforms them.