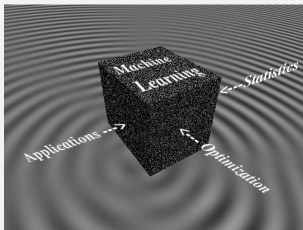


Algorithms for Graph Label Prediction: Challenging Cases

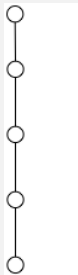
18 November 2008



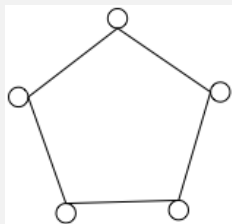
1 Conceptual Cases

2 Practical Challenges

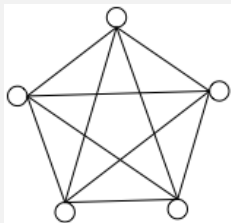
I. Conceptual Cases



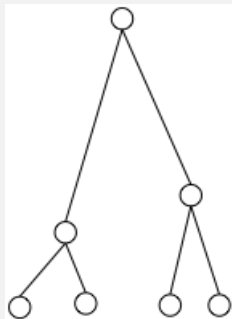
1 Line graph



1 Circle graph



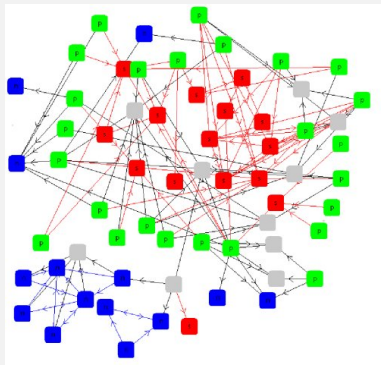
- 1 Clique graph



- 1 Tree graph

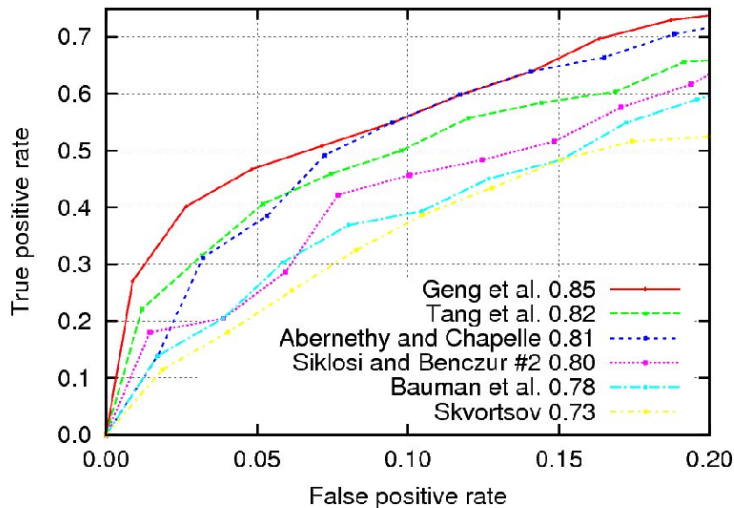
II. Practical Challenges

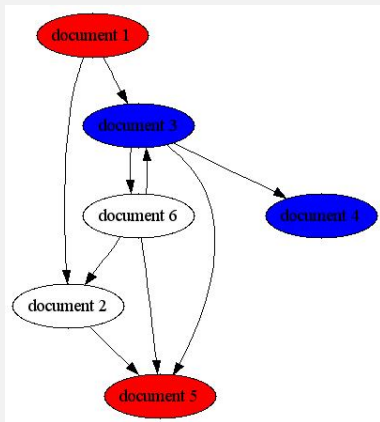
Web Spam Challenge



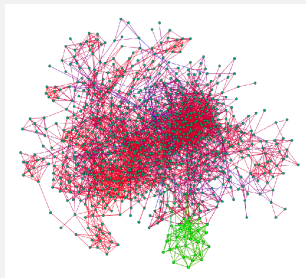
- 1 5,000 Webpages (LIBFR6), 15,000 Host (UK)
- 2 'SPAM', 'NOSPAM', 'BORDERLINE' or 'CANT CLASSIFY'
- 3 Oriented, binary weighted graph
- 4 Additional 'Content-vertex' and 'Content-Edge'
- 5 Evaluation on Subset remaining vertices
- 6 AUC
- 7 <http://webspam.lip6.fr>

Web Spam Challenge

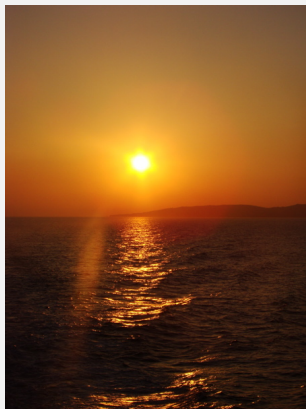




- 1 114,366 documents from the Wikipedia XML Corpus
- 2 10% Labels, 15 Classes
- 3 Classification and Clustering
- 4 <http://xmlmining.lip6.fr>



- 1 Gene-gene interaction networks
- 2 Functional relatedness: "If this gene is relevant to disease X, then most probably also that genes are"



- 1 Benchmark Cases for Online Learning?
- 2 Convert Classification in Graph Labeling?