Link Prediction through Iterative Link Classification

Murali K. Pusala

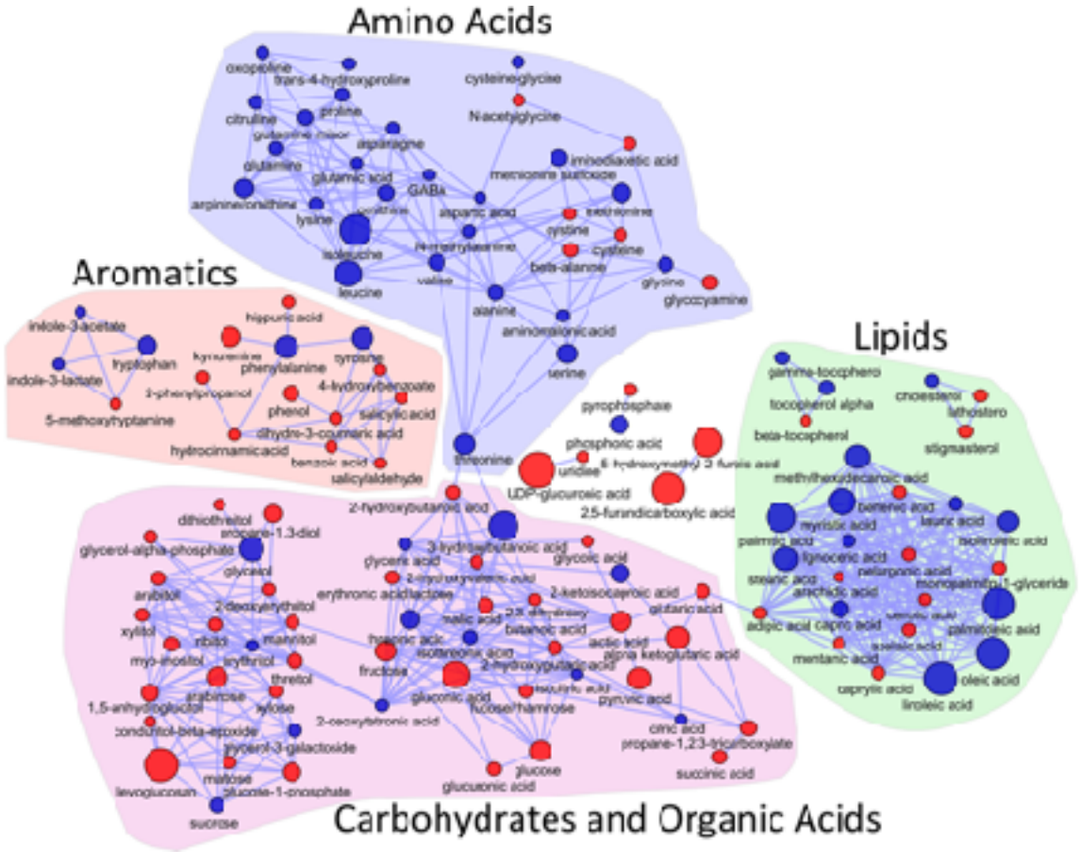
Oct 16 2017

Overview

- **Introduction**
- **Link Prediction**
    - **Definition**
    - **Applications**
    - **Supervised Link Prediction**
- **Iterative Link Classification**
    - **Motivation**
    - **Illustration**
    - **Experiments**
- **Conclusion**
- **Class Projects**

Graphs Everywhere

Social Networks

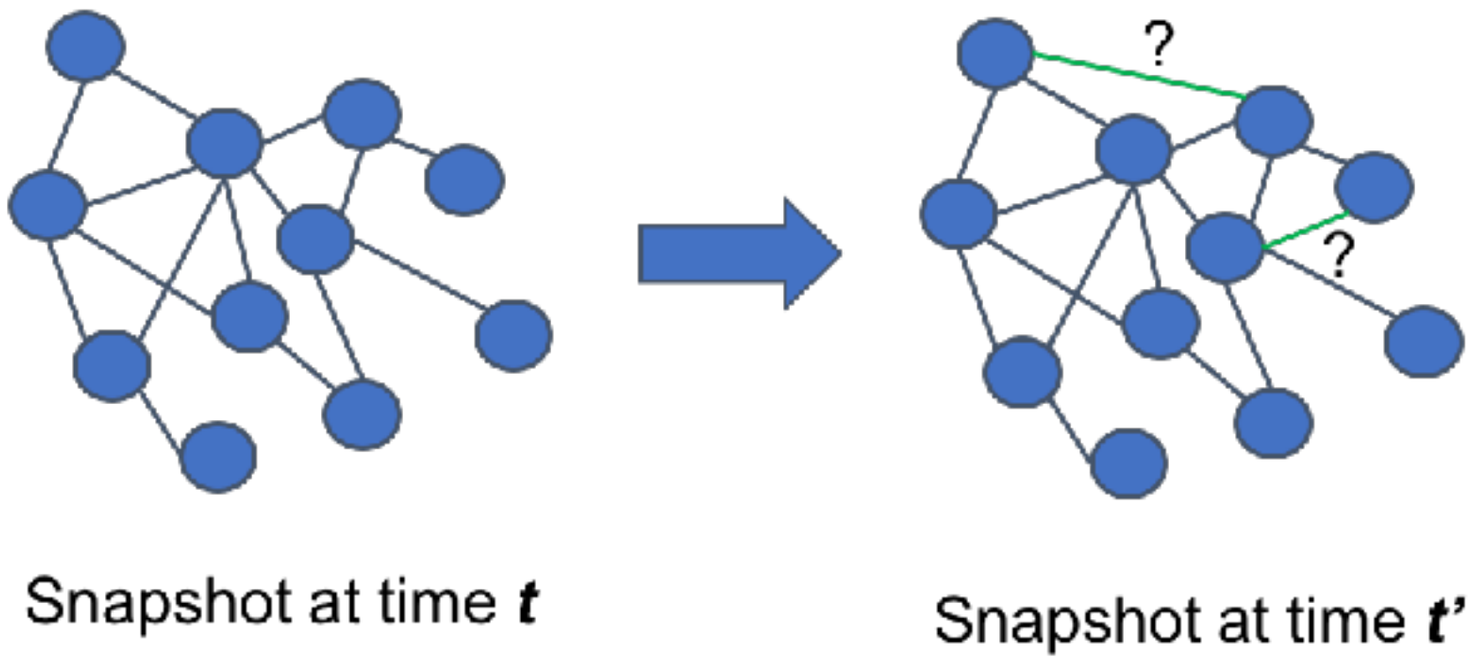


Biological Networks

IoT Networks

Link Prediction

- Given a snapshot of a network at time t, accurately predicting the likelihood of edges that will be added to the network during the interval (t, t′), where t < t′ is referred as Link Prediction

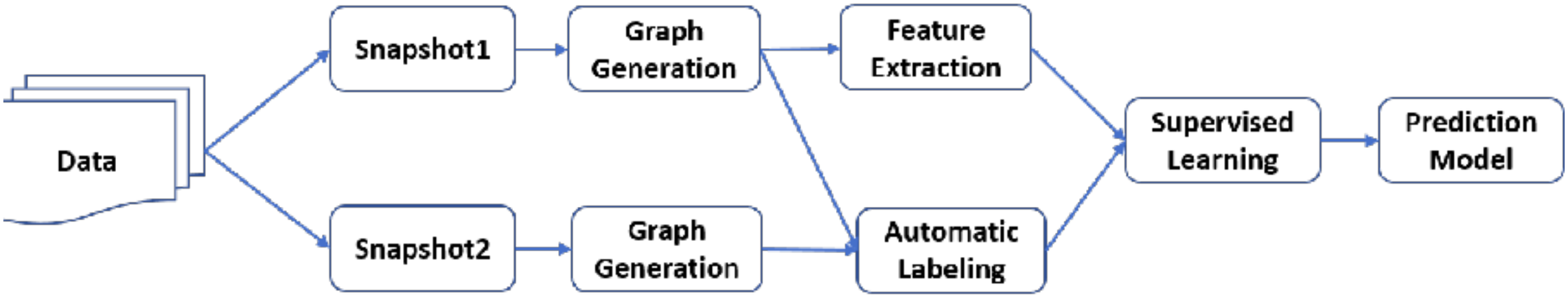


Applications

- Friend suggestions in social networks
- Predict possible interaction between terrorists
- Generating hypothesis from literature
- Collaborative Filtering in recommendation system

Supervised Link Prediction

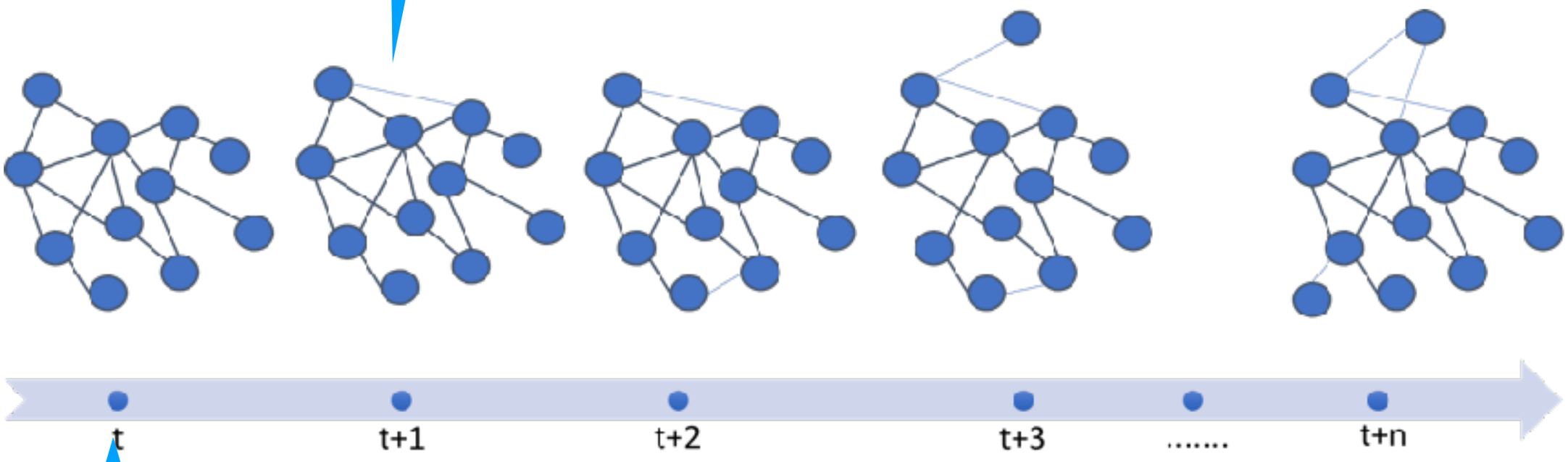
(Katukuri 2012)



Supervised link prediction workflow

Iterative Link Classification: Motivation

Link prediction approaches predicts the network at *t+2* using network at *t+1*

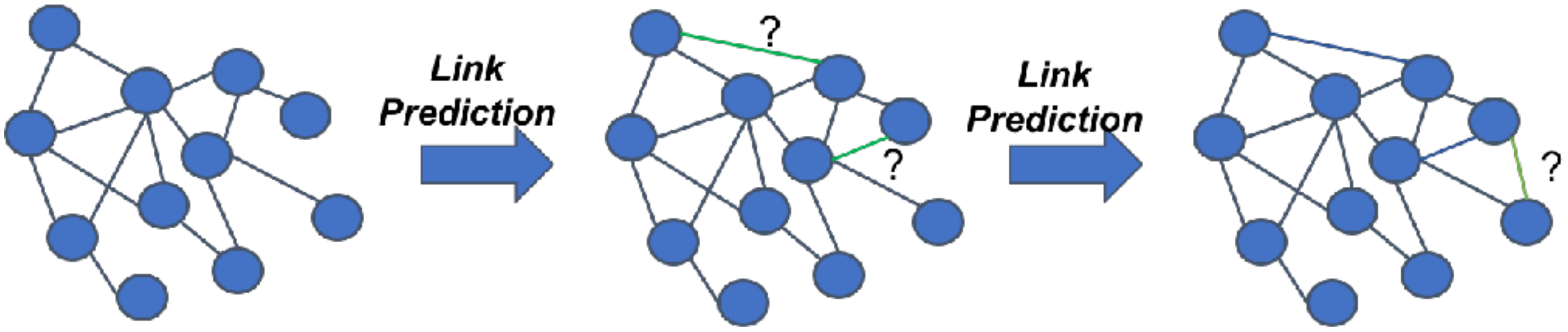


Link prediction approaches predicts the network at *t+1* using network at *t*

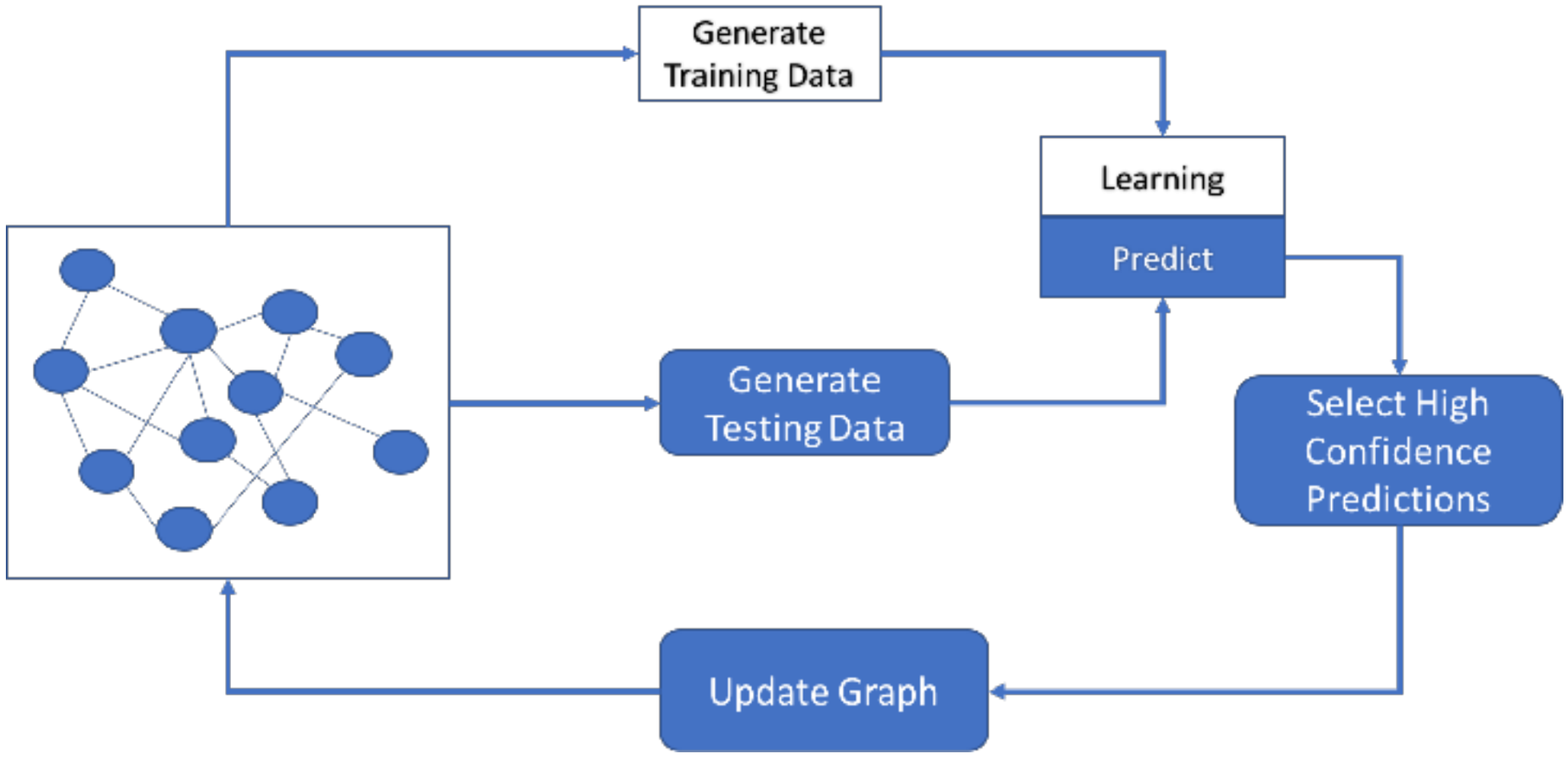
Motivation

- Forecasting in a network presents an unique opportunity to use knowledge to predict links that appear in future

- "Can formation of a relationship, lead to a possible formation of more relationships in the same timeperiod?"



Iterative Link Classification



Iterative Link Prediction Illustration

Experimental Setup

- Experiments designed to study:
  - Effectiveness of the iterative method to predict links
  - Effect of confidence threshold on the performance of the model
  - Effect of disappearing links
  - Effect of machine learning algorithms on the performance of iterative link prediction

Experimental Setup

- Link Prediction as hypothesis generation
- 2014 Medline dataset (1991-2010)
- SVM, C4.5 decision tree
- 5 iterations

Conclusions

- We observed an 6%-7% increase in accuracy over the single pass link prediction and a 10% increase in predicting relevant links compared to traditional link prediction

- Including disappearing links in the prediction approach did not show any improvement in performance the iterative method.

Possible Class Projects

- Evaluate our approach on different domain datasets like social network, citation graph

- Investigating the topological and temporal characteristics of additional links predicted using our iterative method

- Any further question email me at **muralipusala@gmail.com**

- References

  - Katukuri, J. R., Xie, Y., Raghavan, V. V., and Gupta, A. (2012). Hypotheses generation as supervised link discovery with automated class labeling on large-scale biomedical concept networks. *BMC Genomics*, 13(Suppl 3):S5.

  - Lü, Linyuan, and Tao Zhou. "Link prediction in complex networks: A survey." *Physica A: statistical mechanics and its applications* 390.6 (2011): 1150-1170.

  - Neville, J. and Jensen, D. (2000). Iterative classification in relational data. *Learning Statistical Models from Relational Data*, pages 42-49.