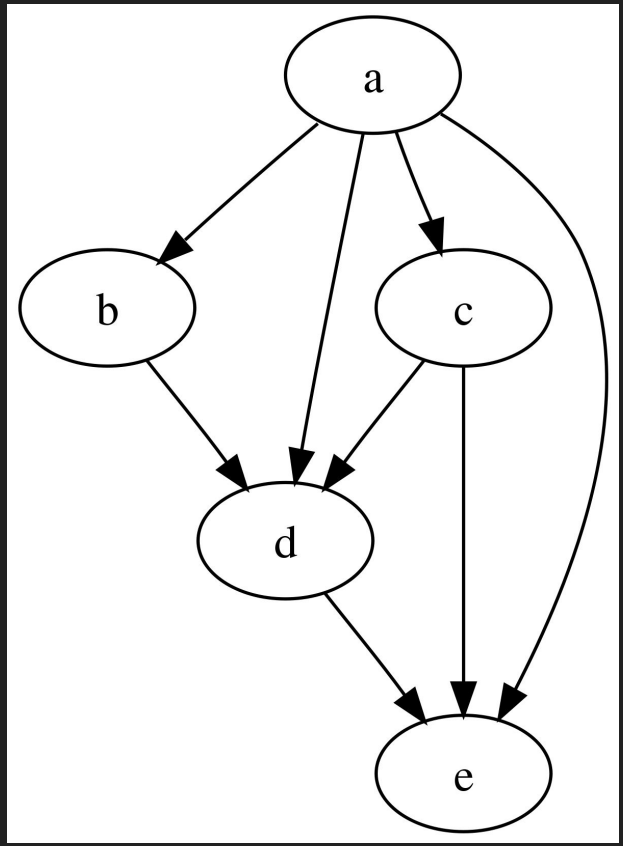
Implementation of Active Learning Methods on Poisson Learning Framework

Faculty Mentor: Professor Jeff Calder
Student researcher: Jason Setiadi

Graphs and Weight Matrix

- Graphs are a useful structure of data
- Research Paper Citation Example
- K-nearest neighbor weight matrix
  - Symmetric
  - Size: # data points x # data points
  - Weight : [0,1]
  - Larger weight = more similarity between the nodes
  - Only keep the weights of k nearest neighbors of each node



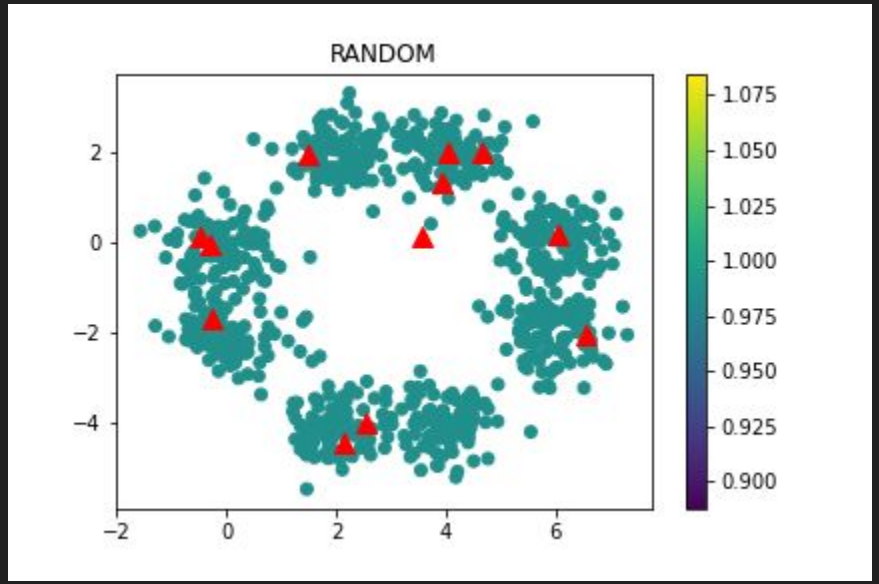
Semi-Supervised Learning

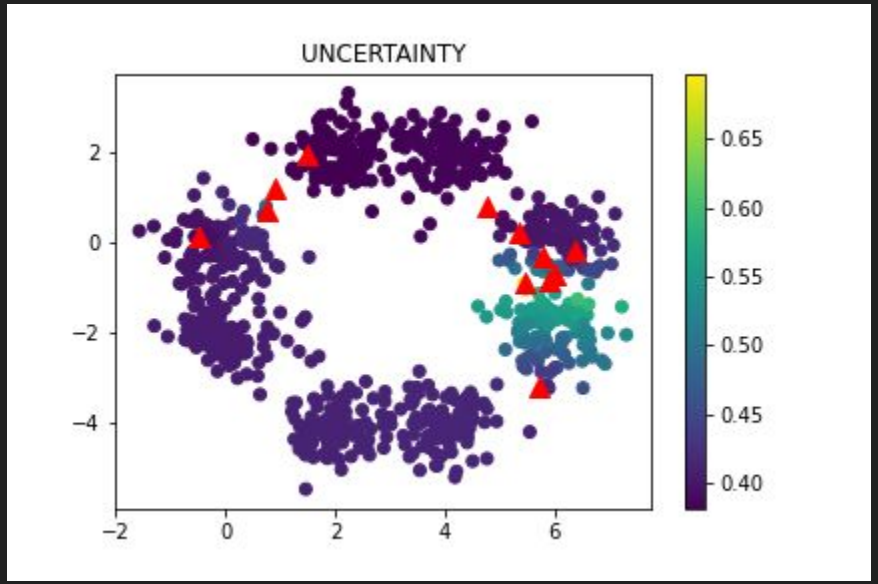
- Training data = data used to train machine learning model
- Labels = output values (numerical, categorical, or ordinal)
- Small amount of training data
- GSSL: Poisson Learning and Laplace Learning

Active Learning

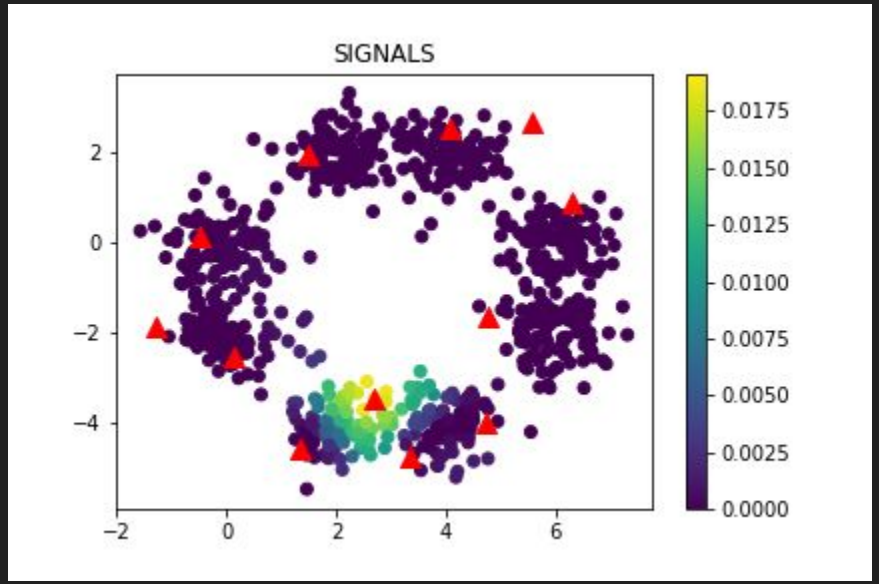
- Method of choosing points to label
- Minimize amount of training data used
- Labels are expensive to obtain
- Three methods of interest:
  - Variance Minimization (V-Optimality)
  - Error Bounds
  - Graph Signals

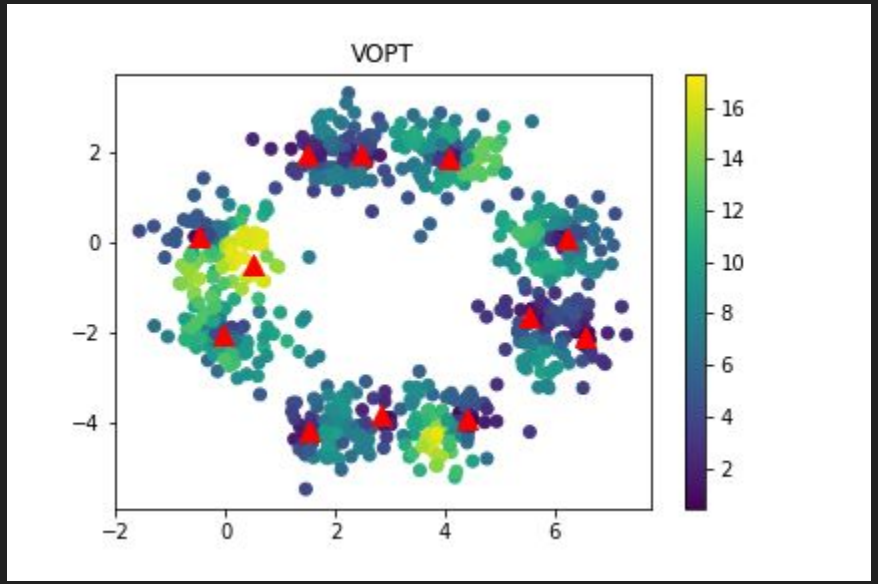
Toy Dataset Example



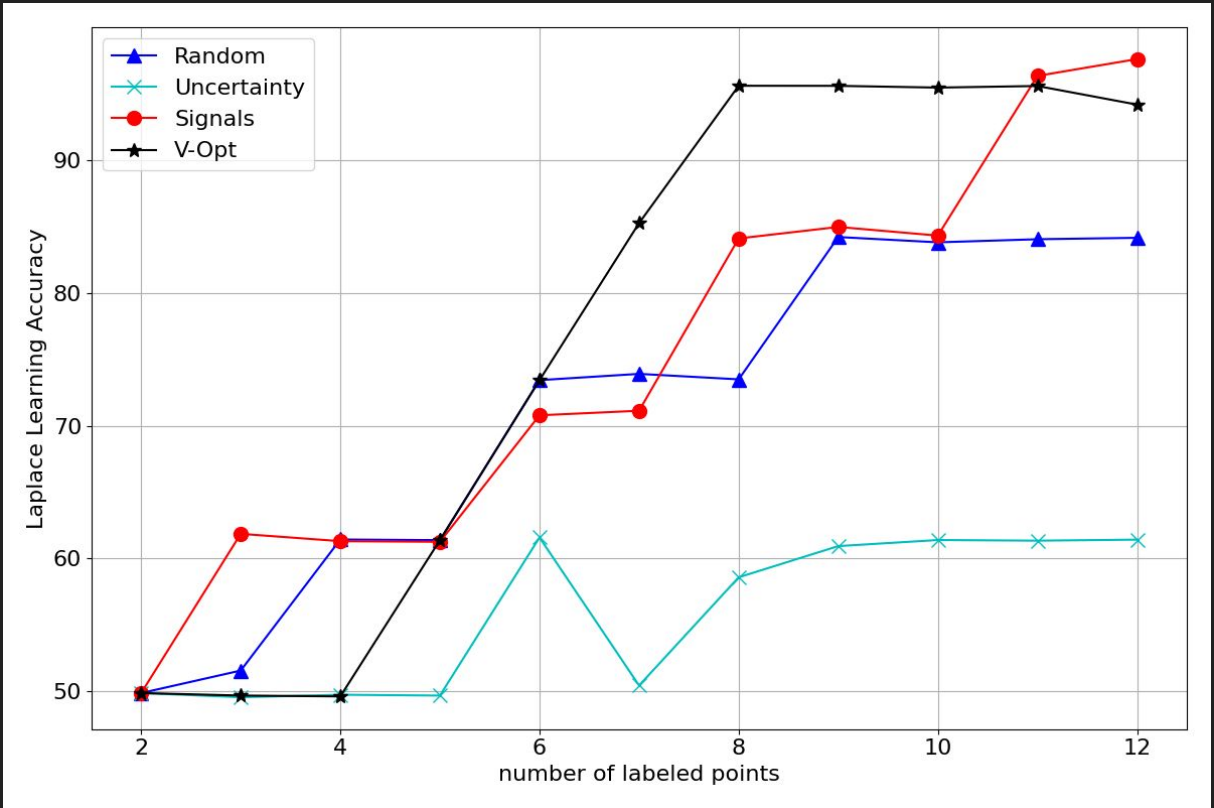


Toy Dataset Example

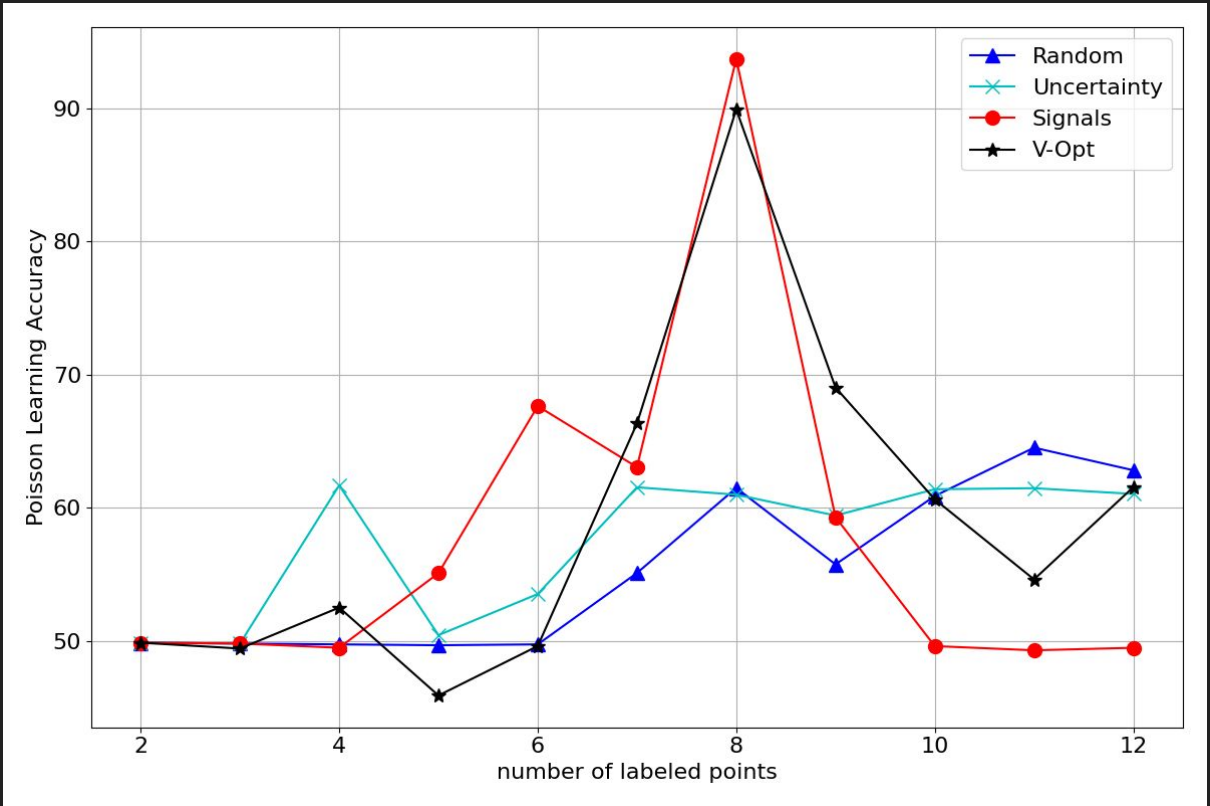




Laplace Learning Results



Poisson Learning Results



Conclusions

- V-Optimality has shown success and is scalable to real-world datasets
- Sampling Theory with Graph Signals is also successful but limited to just small datasets
  - not scalable to real-world datasets because of expensive computation as dataset have more points (need to do eigendecomposition of dense matrix every iteration of active learning)
- Problems with Error Bounds method:
  - Error in mathematical derivation -> Similar to V-Optimality
  - Really large objective function values

Future Work

- Discover why Laplace learning works better on some datasets than Poisson Learning
- Explore other active learning methods that works well with Poisson Learning