AA228 Project 1 Strategy

Somrita Banerjee

19 October 2018

1 Description of algorithm

1.1 Bayesian score function

- The scoring function calculates all the M_{ijk} values when evaluating a graph for the first time
- When adding a parent to node i, it only evaluates the M_{ijk} values for that node i.

1.2 Structure learning search

I implemented a K2 search with the following steps:

1. Start with a prior of a fully unconnected graph
2. Compute score for this graph
3. Try to add a single edge to the graph by going through all the nodes one by one
4. For each node, find all the parents that can be added and add the one that maximizes the Bayesian score. You can't add a node as a parent if:
   - it is the same node as the child
   - it is already a parent of the child
   - adding it causes the graph to become cyclic
5. Return this graph and then re-do the search through all the nodes.

2 Time taken for each graph

- For small.csv, the search returns immediately.
- For medium.csv, the search takes less than 2 seconds (without printing) to converge.
- For large.csv, the search takes ≈250 steps to converge and takes 1 hour.

3 Graph Plots

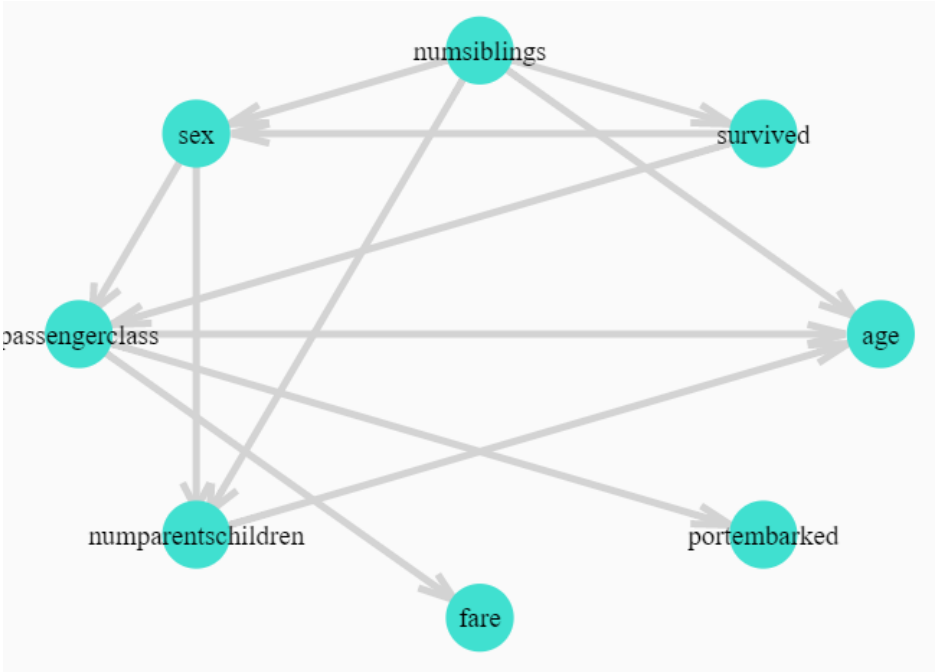

Figure 1: Graph for small.csv

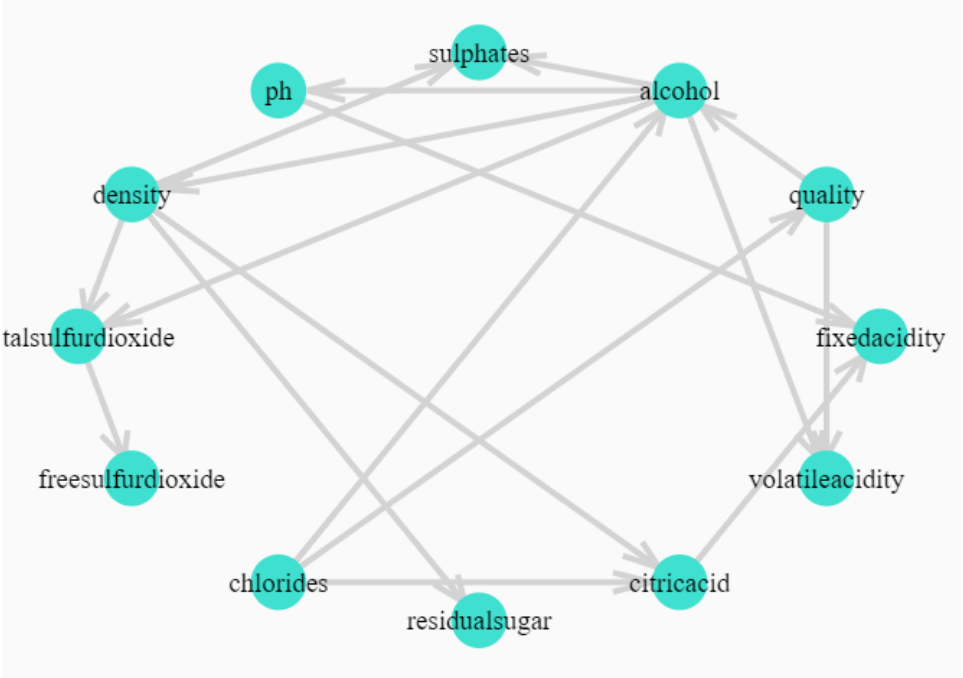

Figure 2: Graph for medium.csv

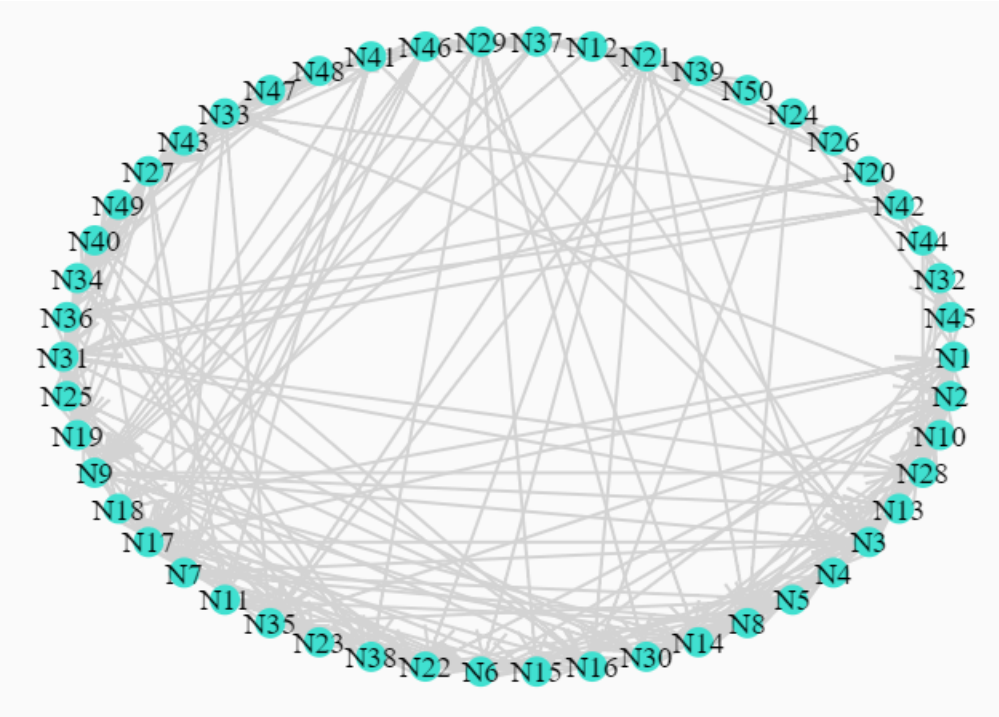

Figure 3: Graph for large.csv