**INFSCI 2725: Data Analytics**

**Assignment 3: Validation and Testing**

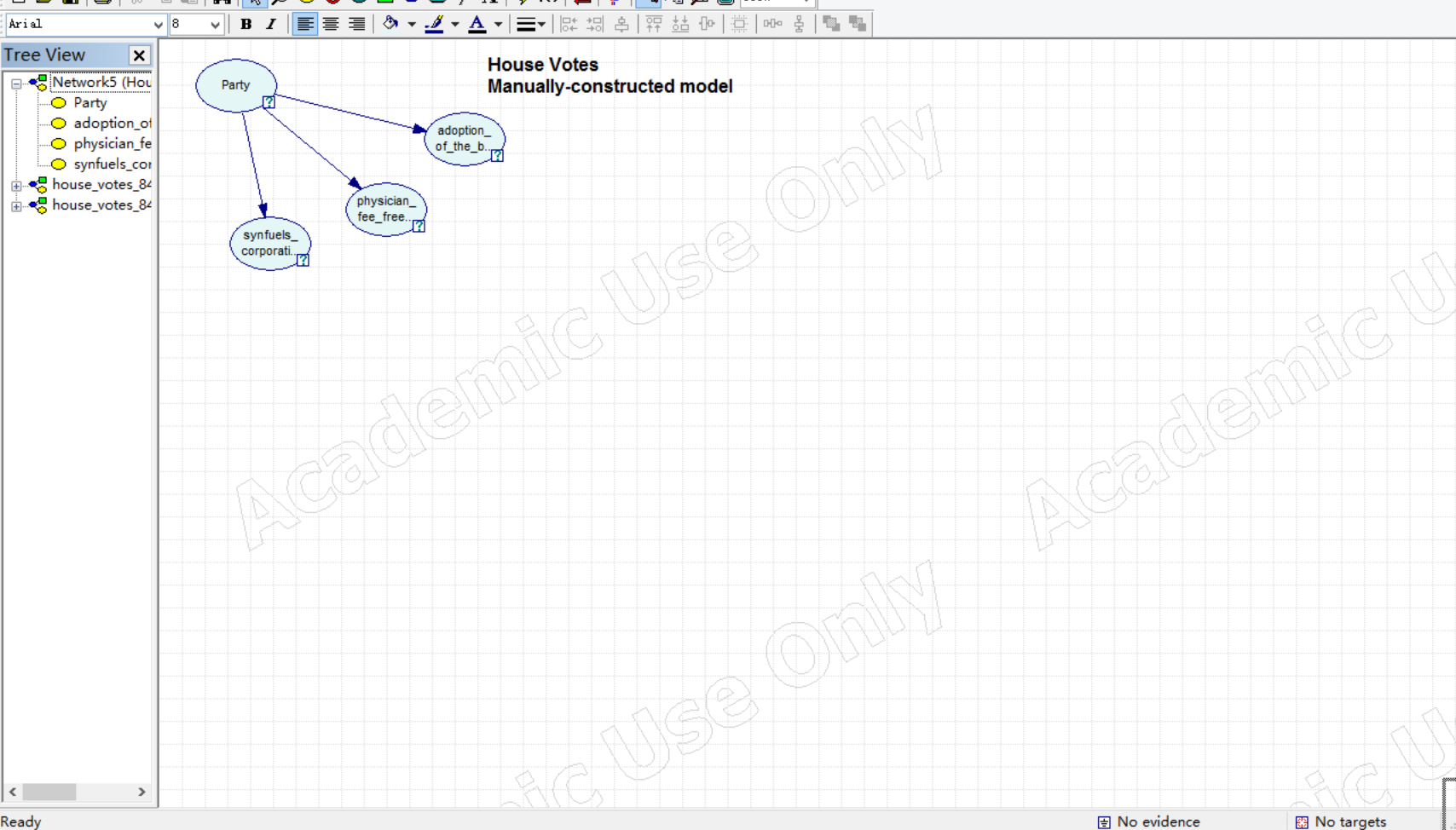
**Team members**: Yumeng Lu

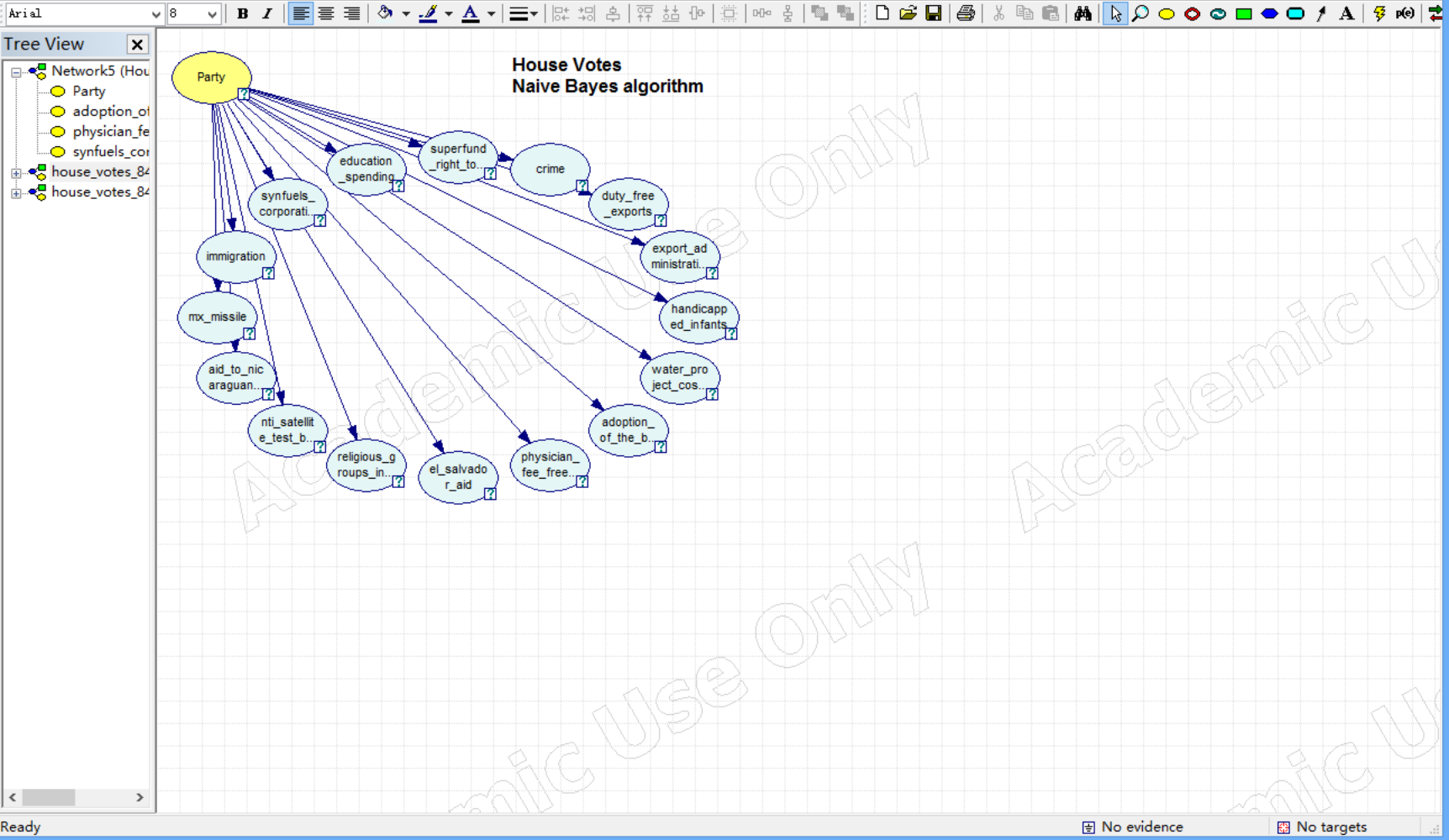
Zhaoxuan Ren

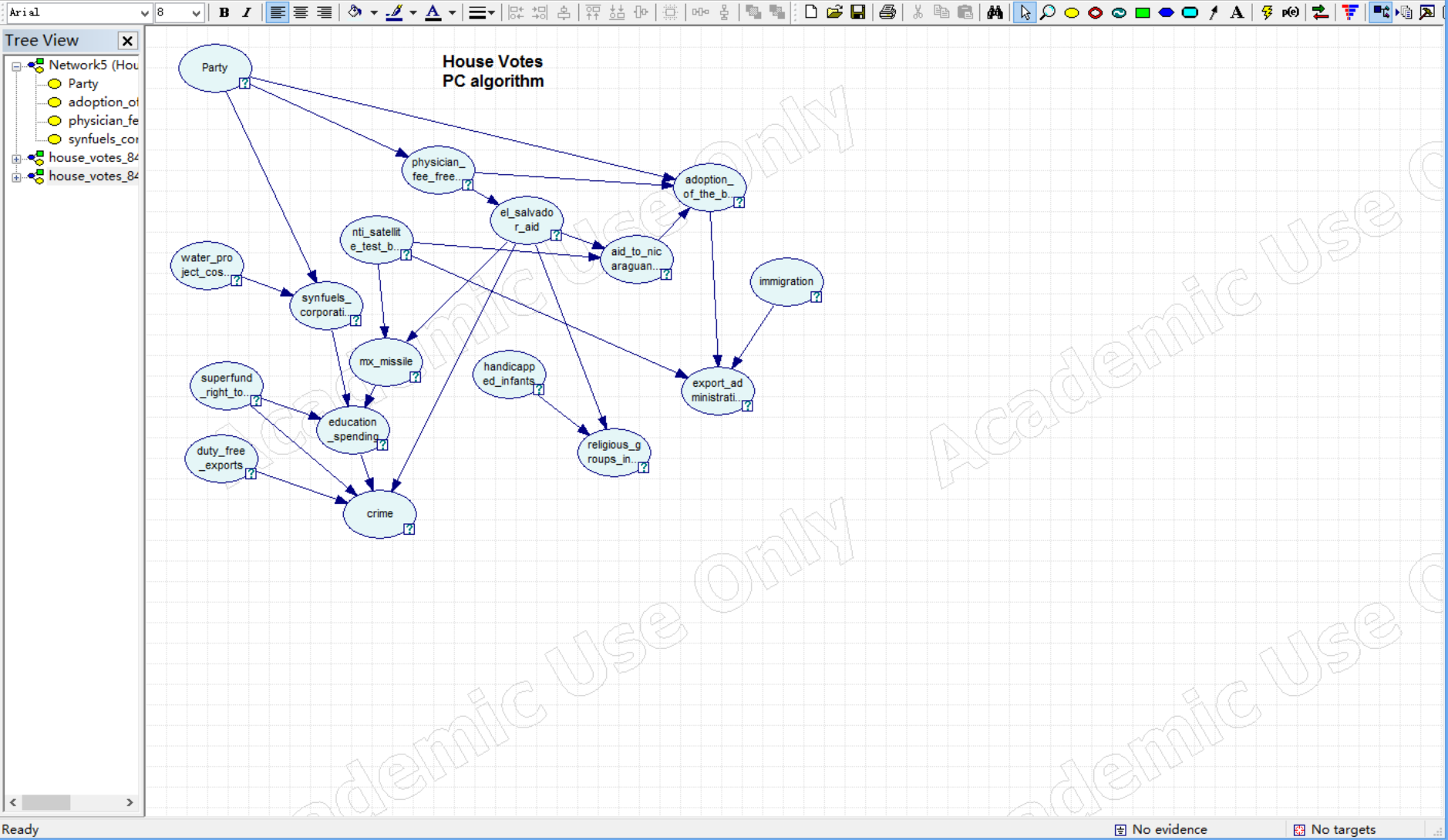
Tao Li

In assignment 3, we use GeNIe to the tasks.

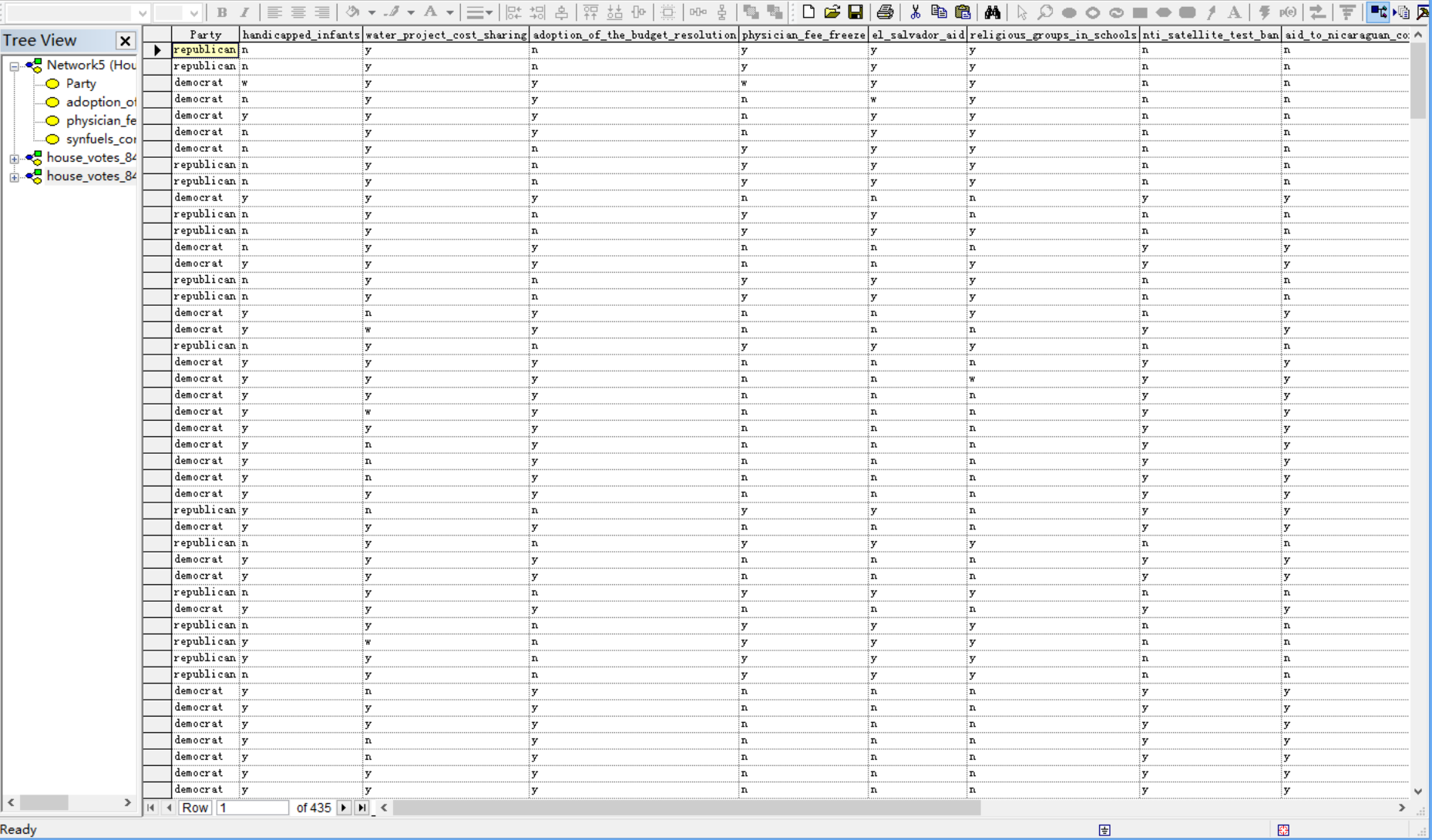
Firstly, input the GeNIe models:



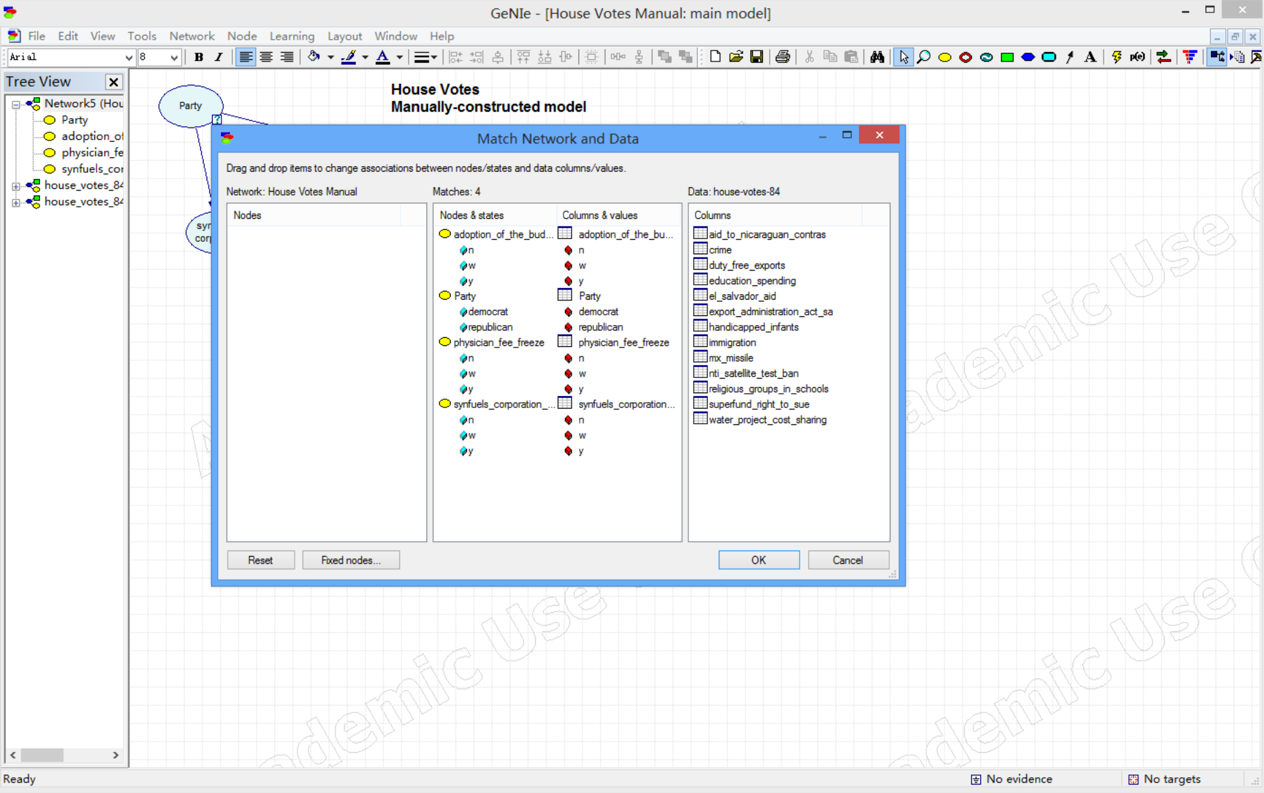




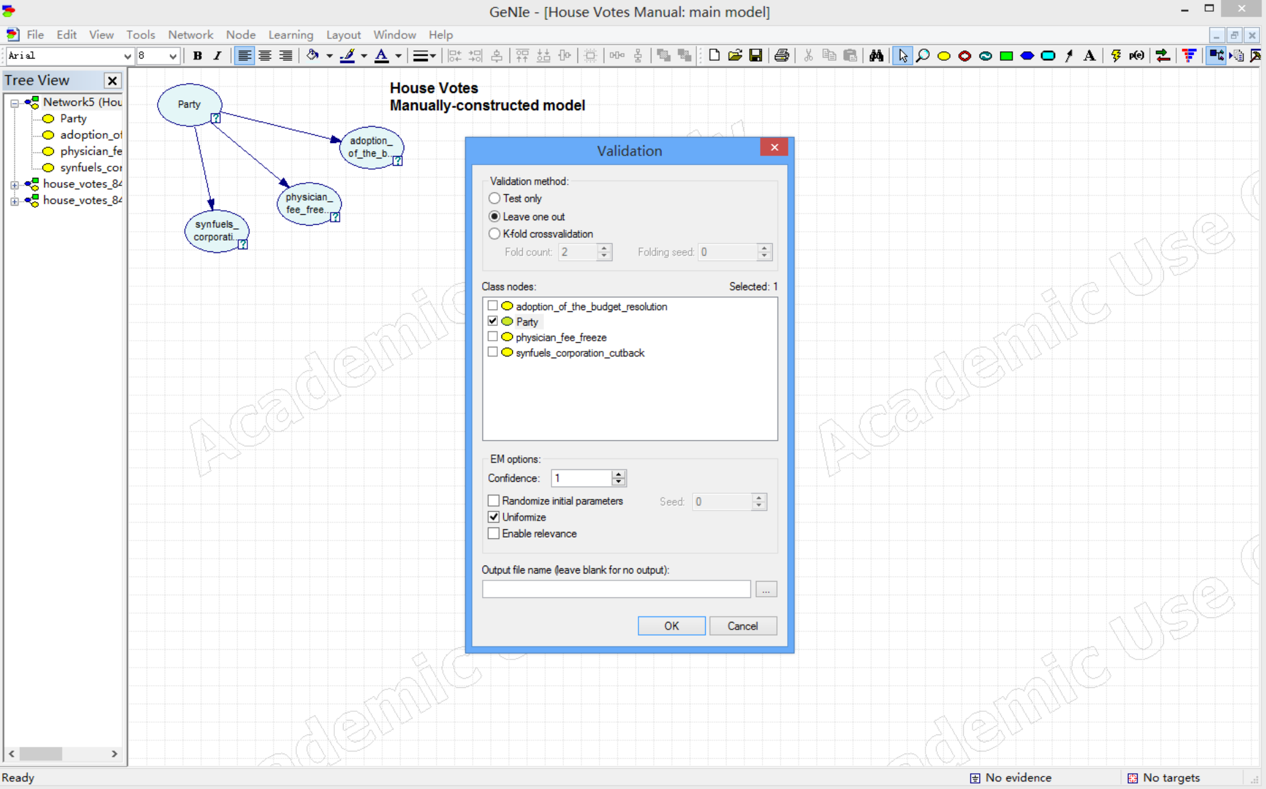
The import the data file:



Use validation to test the model:



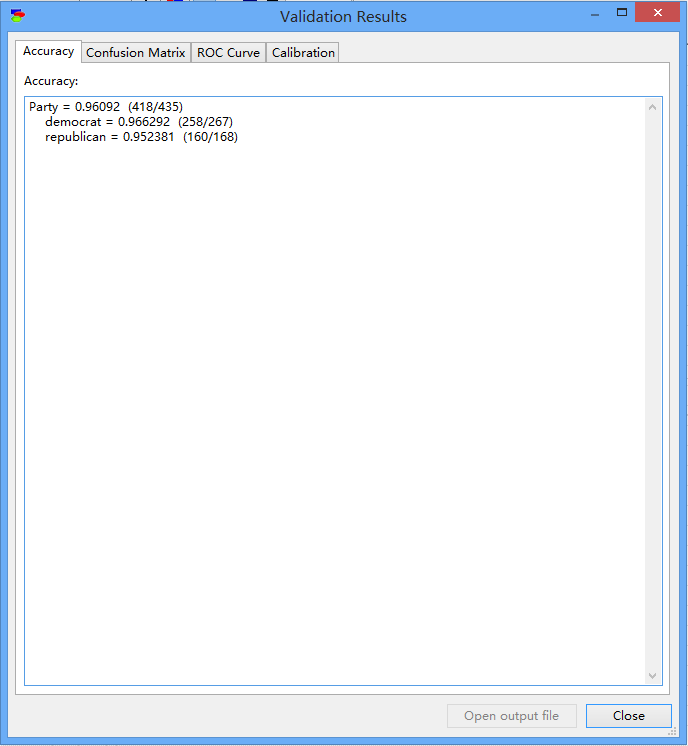
Choose validation method as leave-one-out, and the class nodes to be Party:



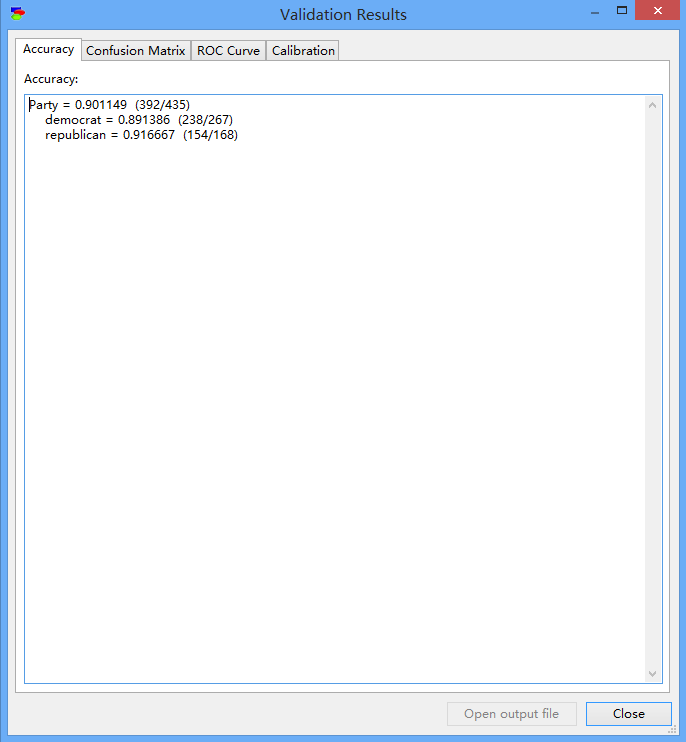
**Tasks:**

1. **Overall classification accuracy**

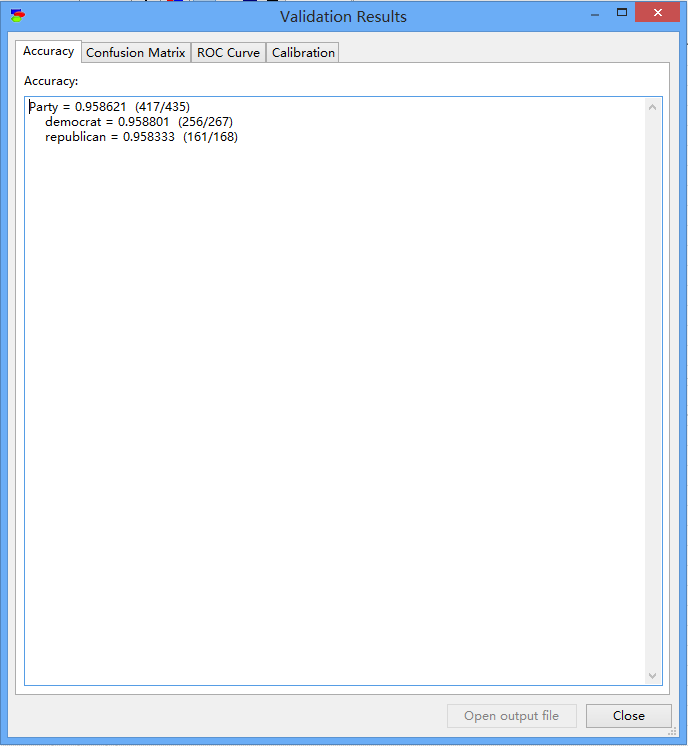
Manually-constructed model:



Naive Bayes algorithm:



PC algorithm:



1. **Sensitivity and specificity for each of the two parties**

**House Votes Manual.xdsl**

|  |  |  |
| --- | --- | --- |
|  | democrat | republican |
| democrat | 258 | 9 |
| republican | 8 | 160 |

Sensitivity=258/(258+8)=0.96992481

Specificity=160/(160+9)=0.94674556

**House Votes Naive.xdsl**

|  |  |  |
| --- | --- | --- |
|  | democrat | republican |
| democrat | 238 | 29 |
| republican | 14 | 154 |

Sensitivity=238/(238+14)=0.9444444444

Specificity=154/(154+29)=0.8415300546

**House Votes PC.xdsl**

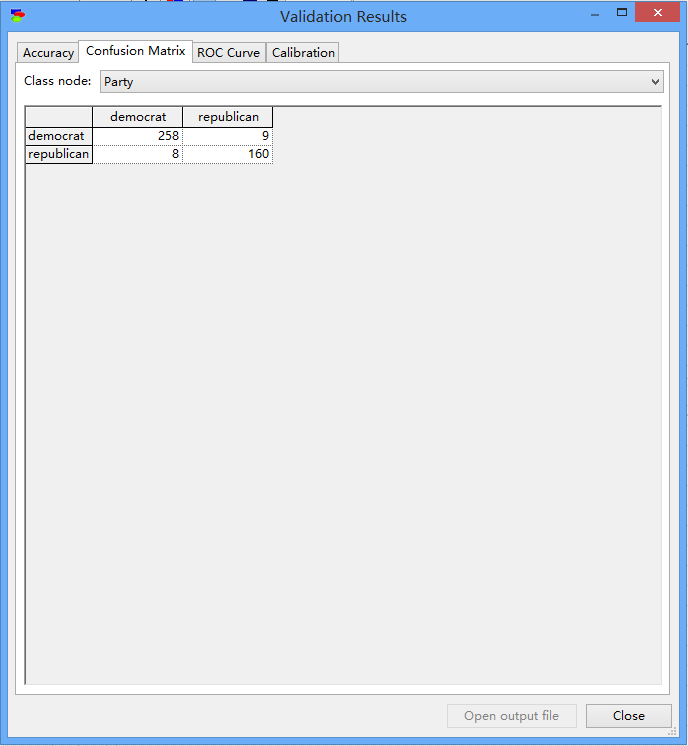
|  |  |  |
| --- | --- | --- |
|  | democrat | republican |
| democrat | 258 | 9 |
| republican | 5 | 163 |

Sensitivity=258/(258+5)=0.980988

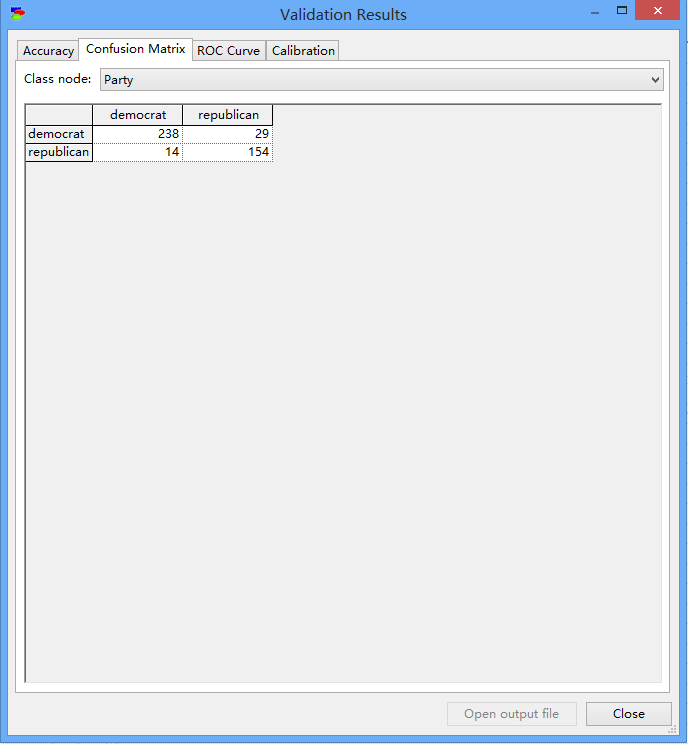
Specificity=163/(163+9)=0.947674

1. **Positive and negative predictive value for each of the two parties**

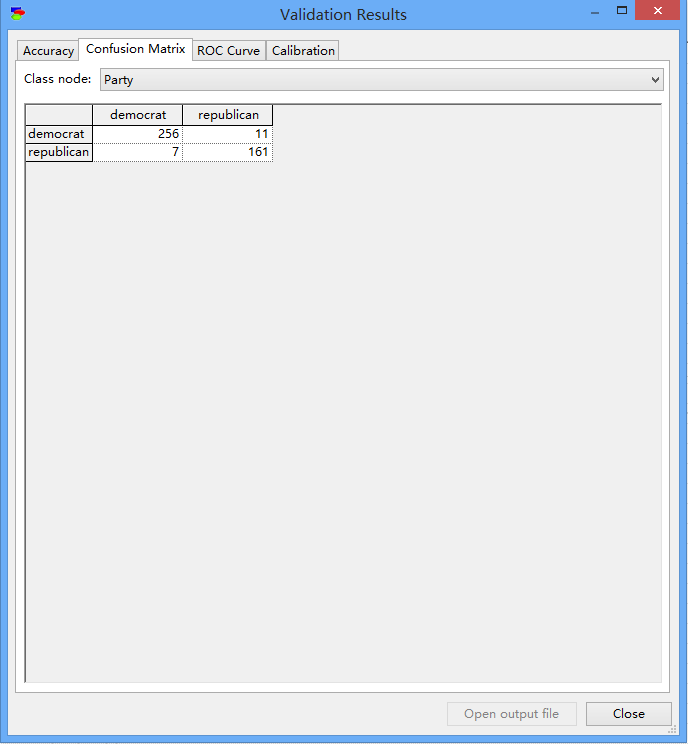
Manually-constructed model:



Naive Bayes algorithm:

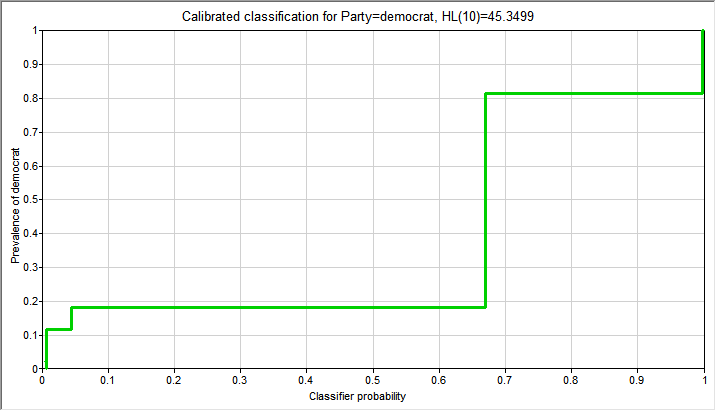


PC algorithm:

****

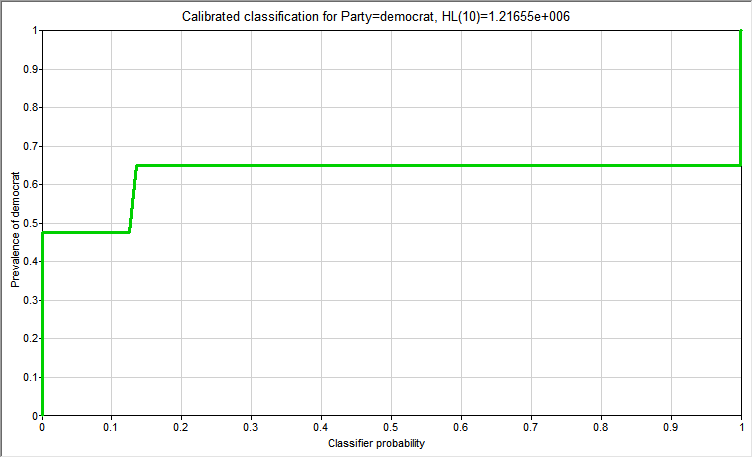
1. **Calibration curve for a selected bin count or window size**

1)House Vote Manual.xdsl



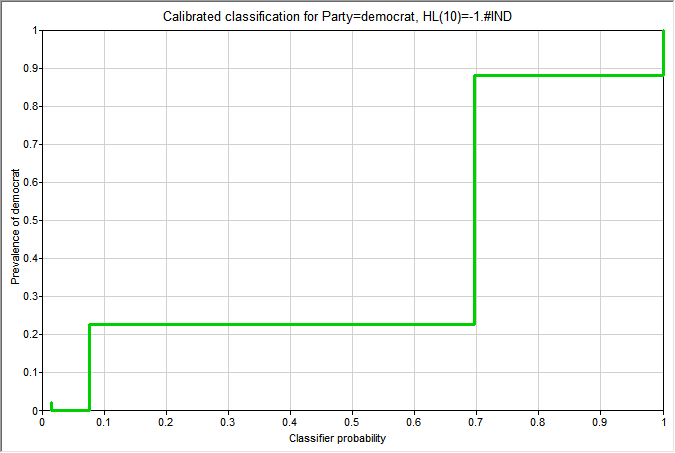
Bin count:10

2) House Votes Naive.xdsl



Bin count:10

3) House Votes PC.xdsl



Bin count: 10