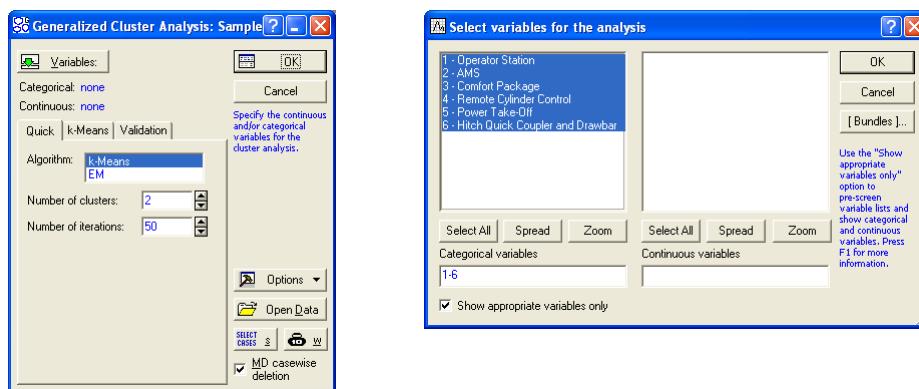


Outline

- Clustering
- Association rules
- Graphic user interface

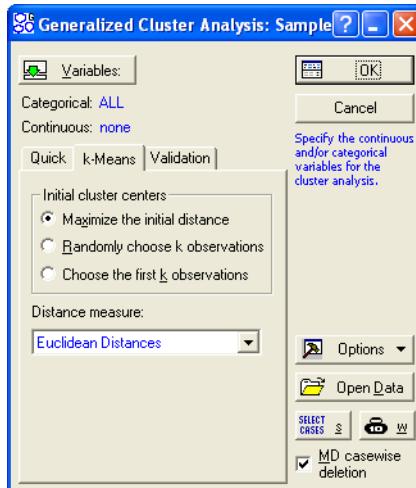
Clustering

- Select k-Means and choose the variables



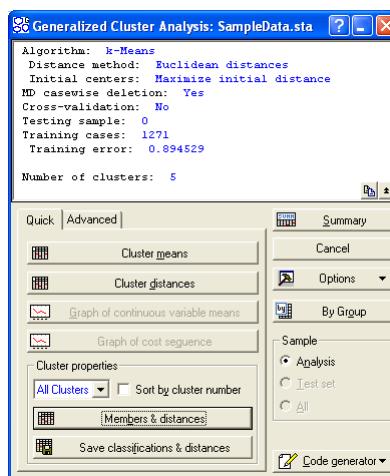
Clustering

- Choose the distance metrics and initial cluster centers



Clustering

- 5 clusters and see the results



Clustering

- Centroids (cluster means)

Centroids for k-means clustering (SampleData.sta)								
	Operator	Station	AMS	Comfort Package	Remote Cylinder Control	Power Take-Off	Hitch Quick Coupler and Drawbar	Number of cases
Cluster								Percentage(%)
1		991	1005	1210	2400	2410	2530	386 30.36979
2		989	1005	1200	2400	2410	2530	420 33.04485
3		991	1005	1200	2400	2410	2530	150 11.80173
4		986	1005	1200	2300	2410	2530	152 11.95909
5		989	1005	1200	2300	2420	2530	163 12.82455

Clustering

- Members and their distance to the centroids

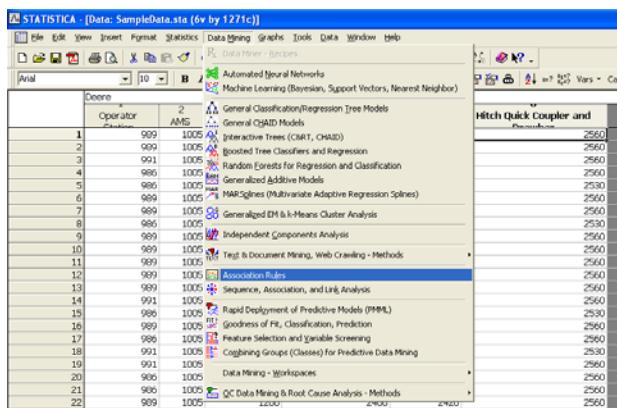
Cluster members (SampleData.sta)								
	Case No.	Final classification	Operator	Station	AMS	Comfort Package	Remote Cylinder Control	Power Take-Off
	Case No.							Hitch Quick Coupler and Drawbar
67		3	991	1005	1200	2400	2410	2560 1.000000
68		2	989	1005	1200	2400	2410	2560 1.000000
69		2	989	1005	1200	2400	2410	2560 1.000000
70		2	989	1005	1200	2400	2410	2560 1.000000
71		2	989	1005	1200	2400	2410	2560 1.000000
72		2	989	1005	1200	2400	2410	2560 1.000000
73		2	989	1005	1200	2300	2410	2530 1.000000
74		2	989	1005	1200	2300	2410	2530 1.000000
75		2	989	1005	1200	2400	2410	2560 1.000000
76		5	989	1005	1200	2300	2420	2530 0.000000
77		5	989	1005	1200	2300	2420	2530 0.000000
78		2	989	1005	1200	2400	2410	2560 1.000000
79		2	989	1005	1200	2400	2420	2560 1.414214
80		2	989	1005	1200	2400	2410	2530 0.000000
81		2	989	1005	1200	2400	2410	2560 1.000000

Centroids for k-means clustering (SampleData.sta) Cluster members (SampleData.sta)

A... SampleData.sta C1,V1 1 CAP NUM IR

Association rules

- Use the Deere data set



Association rules

- Select variables and set up proper parameters

The image shows two overlapping dialog boxes. The top dialog is titled "Select categorical variables" and lists variables: 1 - Operator Station, 2 - AMS, 3 - Comfort Package, 4 - Remote Cylinder Control, 5 - Power Take-Off, and 6 - Hitch Quick Coupler and Drawbar. It has buttons for OK, Cancel, and [Bundles...]. Below it is a message: "Use the 'Show appropriate variables only' option to pre-screen variables and show categorical and continuous variables only." The bottom dialog is titled "Association Rules: SampleData.sta" and has tabs for Quick and Advanced. Under "Options for results", it shows: Minimum support: .45, Minimum confidence: .44, Minimum correlation: .50, Maximum item set size in body: 10, and Maximum item set size in head: 10. It also has buttons for OK, Cancel, Options, Open Data, SELECT CASES, and Weighted moments.

Association rules

- See rules

Summary of association rules (SampleData.sta)						
Min. support = 40.0%, Min. confidence = 44.0%, Min. correlation = 50.0%						
Max. size of body = 10, Max. size of head = 10						
Body		=>		Head		Support(%)
1	Comfort Package == 1200	=>		Power Take-Off == 2410	46.57750	69.81132
2	Comfort Package == 1200	=>		Hitch Quick Coupler and Drawbar == 2530	44.92526	67.33491
3	Remote Cylinder Control == 2300	=>		Hitch Quick Coupler and Drawbar == 2530	42.64359	83.00153
4	Power Take-Off == 2410	=>		Comfort Package == 1200	46.57750	74.37186
5	Hitch Quick Coupler and Drawbar == 2530	=>		Comfort Package == 1200	44.92526	68.54742
6	Hitch Quick Coupler and Drawbar == 2530	=>		Remote Cylinder Control == 2300	42.64359	73.48864
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						

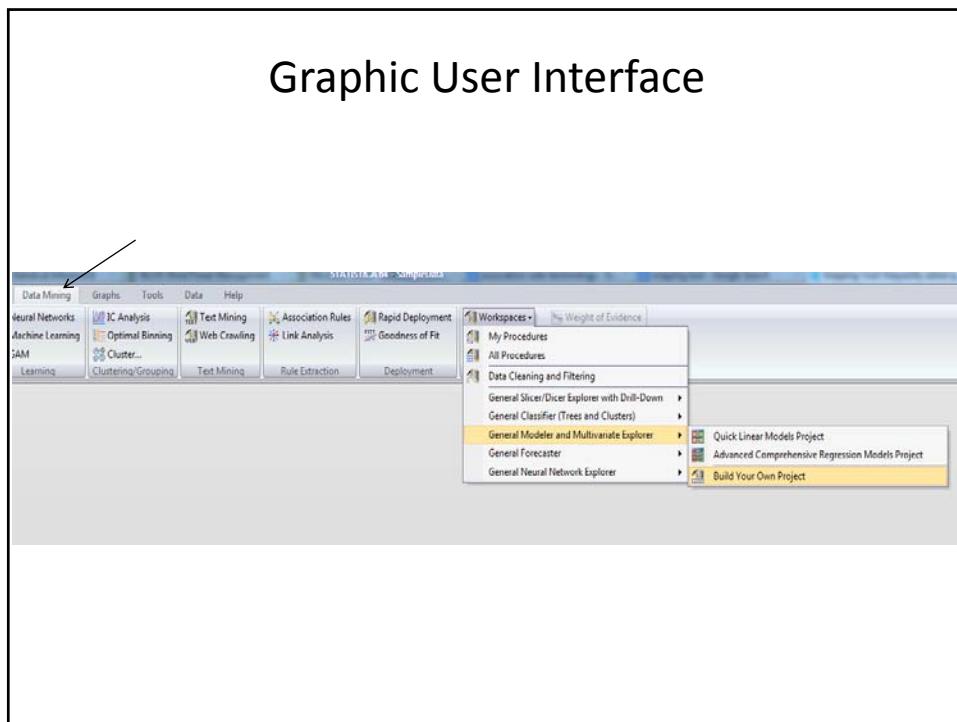
Graphic User Interface

- Divide CPU data into training and testing data set

Data: CPU_Train.sta (7r by 19c)						
1	NWYC1	2	MMIN	MMAX	4	CACH
5	CHMIN	6	CHMAX	7	Performan	
1	128	256	512	1024	16	128
2	20	8000	32000	32	6	32
3	29	8000	32000	32	8	32
4	29	8000	32000	32	8	32
5	29	8000	16000	32	8	16
6	26	8000	32000	64	8	32
7	29	8000	32000	48	16	32
8	29	16000	32000	64	16	32
9	23	16000	64000	64	16	32
10	23	32000	64000	128	32	64
11	400	1000	3000	1	2	2
12	400	512	3500	4	1	6
13	60	2000	8000	65	1	8
14	50	4000	16000	65	1	8
15	29	64	64	1	4	4
16	200	512	16000	0	4	32
17	167	512	2000	0	4	15
18	143	512	5000	7	4	29
19	143	1000	3000	0	2	16
20	110	5000	5000	142	8	64
21	143	1500	6300	0	5	32
22	143	3100	6200	0	5	20
23	143	3100	6200	0	6	64
24	110	3100	6200	0	6	64
25	320	128	6000	0	1	12
26	320	512	2000	4	1	21
27	320	256	3000	4	1	6
28	320	256	3000	4	1	3
29	320	512	5000	4	1	5

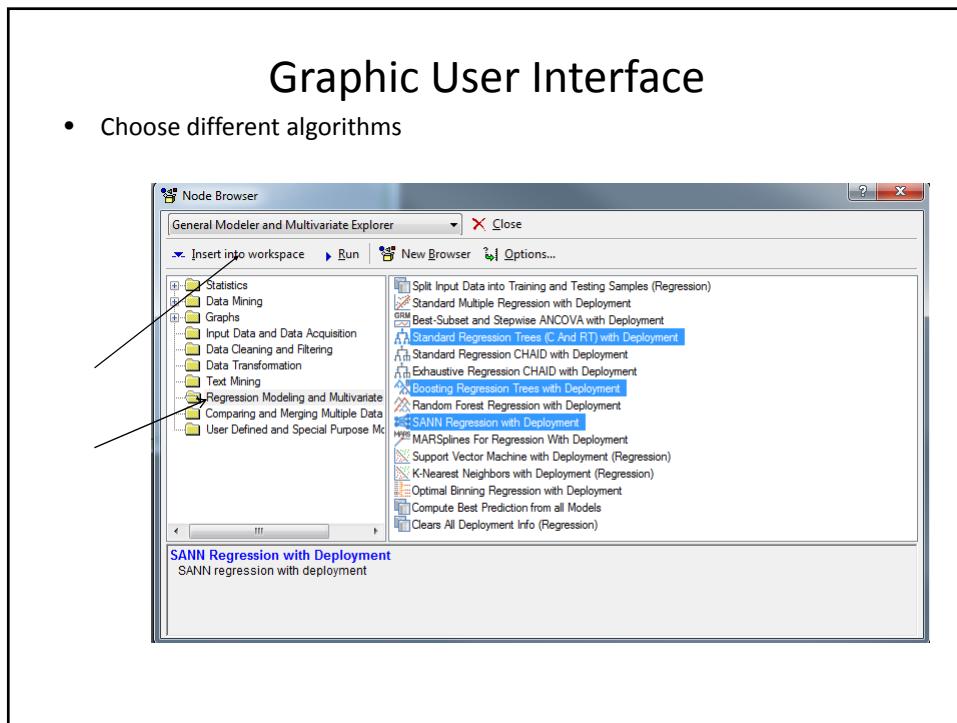
Data: CPU_Test.sta (7r by 19c)						
1	NWYC1	2	MMIN	MMAX	4	CACH
5	CHMIN	6	CHMAX	7	Performan	
1	503	1000	4000	32000	112	52
2	30	8000	64000	96	12	176
3	30	8000	64000	96	12	176
4	180	256	4000	0	1	3
5	180	512	4000	0	1	3
6	180	252	4000	0	1	3
7	180	300	4000	0	1	3
8	124	1000	9000	0	1	8
9	98	1000	8000	32	2	8
10	125	2000	8000	0	2	34
11	400	512	8000	32	0	0
12	400	1000	4000	0	0	0

Graphic User Interface



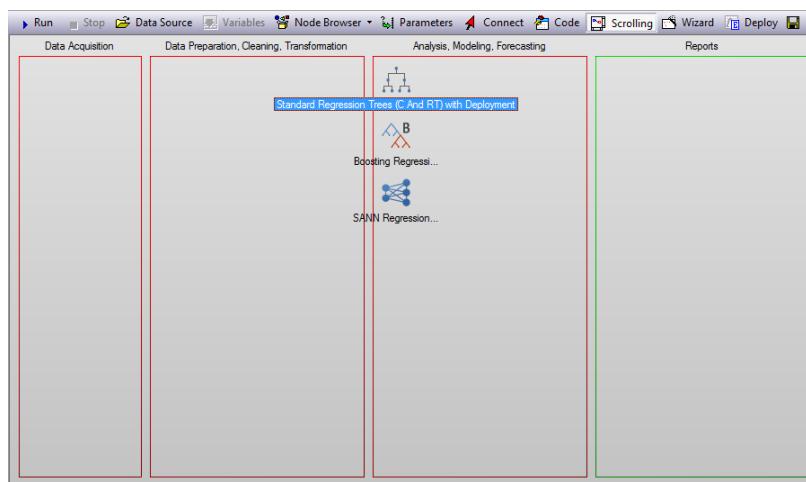
Graphic User Interface

- Choose different algorithms



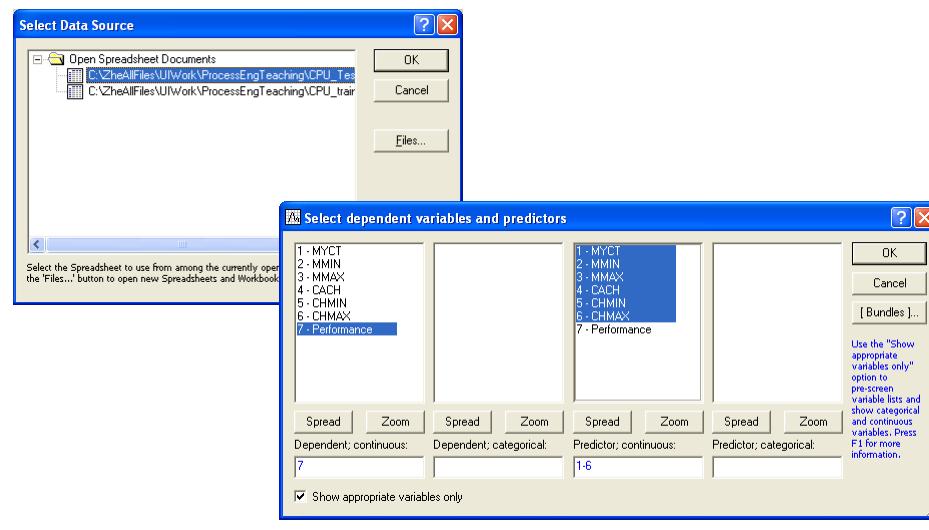
Graphic User Interface

- Insert the selected data mining algorithms into workspace



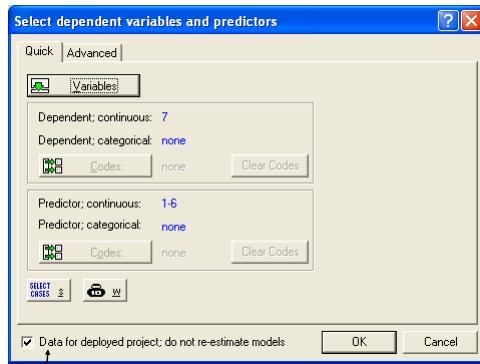
Graphic User Interface

- Select data sources



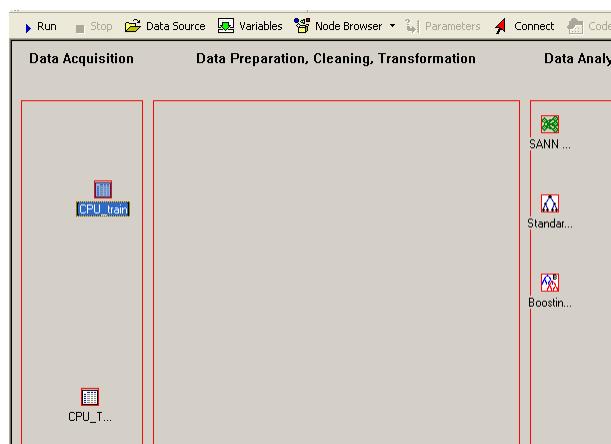
Graphic User Interface

- Specify whether the data is used to build the model or used as a testing set



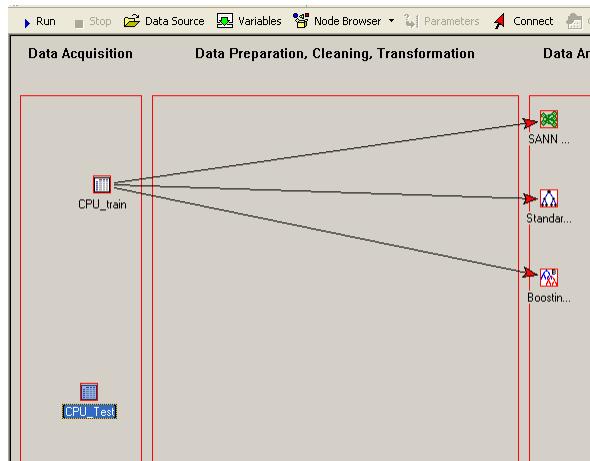
Graphic User Interface

- Connect the data with data mining algorithms



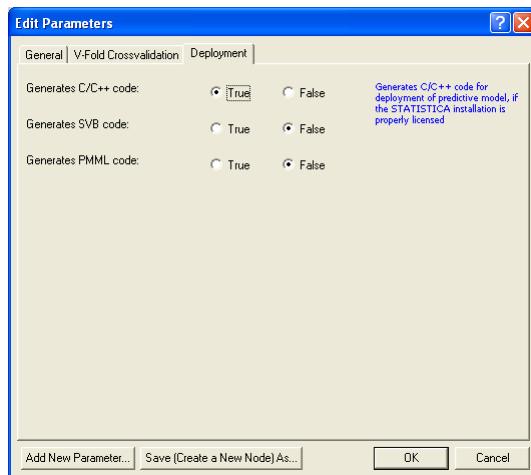
Graphic User Interface

- Connect the data with data mining algorithms



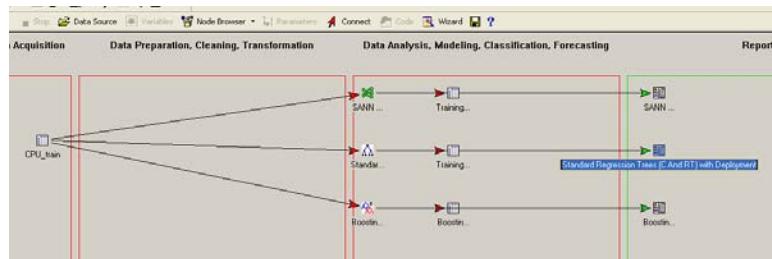
Graphic User Interface

- Set up deployment, double click the data mining algorithm icon



Graphic User Interface

- Click “Run” button



Graphic User Interface

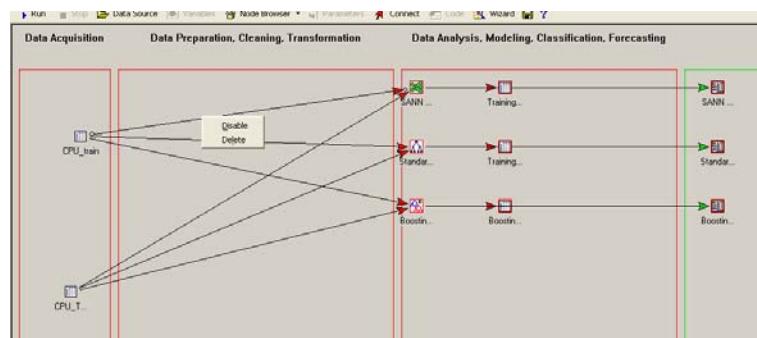
- See the deployment code by double click the icons in “Reports” section

Graphic User Interface

- Test the learnt models by testing data set
- First disable the connections between training data set and the data mining algorithms
- Connect the testing data set with the data mining algorithms

Graphic User Interface

- Test the learnt models



Graphic User Interface

- See the prediction results

