Diabetes Health Indicators

Obs	Diabetes	HighBP	HighChol	CholCheck	Smoker	Stroke	HeartDiseaseorAttack	PhysActivity	Fruits	Veggies	HvyAlcoholConsump	AnyHealthcare
1	0	0	0	1	1	0	0	0	1	1	0	1
2	2	1	1	1	0	0	1	0	1	0	0	1
3	2	1	0	1	0	0	0	1	1	1	0	1
4	2	0	0	1	0	1	1	1	1	1	0	1
5	0	0	1	1	1	0	0	0	0	0	0	1
6	0	1	1	1	1	0	0	1	1	1	0	1
7	0	1	0	1	1	0	0	0	0	1	0	1
8	0	0	0	1	1	0	0	1	0	1	0	1
9	2	0	1	1	0	0	1	0	0	1	0	1
10	0	1	0	1	0	0	0	0	1	1	0	1
11	0	1	0	1	1	0	0	0	0	0	0	1
12	2	0	1	1	1	0	0	0	1	0	1	0	1
13	2	1	0	1	0	0	0	0	1	1	0	1
14	0	0	0	0	0	0	0	0	1	1	1	0	1
15	2	0	1	1	0	0	0	0	0	1	0	1

Diabetes Health Indicators

The HPSPLIT Procedure

Performance Information	
Execution Mode	Single-Machine
Number of Threads	2

Data Access Information			
Data	Engine	Role	Path
WORK.DIABETES_2021	V9	Input	On Client

Model Information	
Split Criterion Used	Entropy
Pruning Method	Cost-Complexity
Subtree Evaluation Criterion	Cost-Complexity
Number of Branches	2
Maximum Tree Depth Requested	10
Maximum Tree Depth Achieved	10
Tree Depth	10
Number of Leaves Before Pruning	893
Number of Leaves After Pruning	219

Number of Observations Read	267124
Number of Observations Used	267124

Diabetes Health Indicators
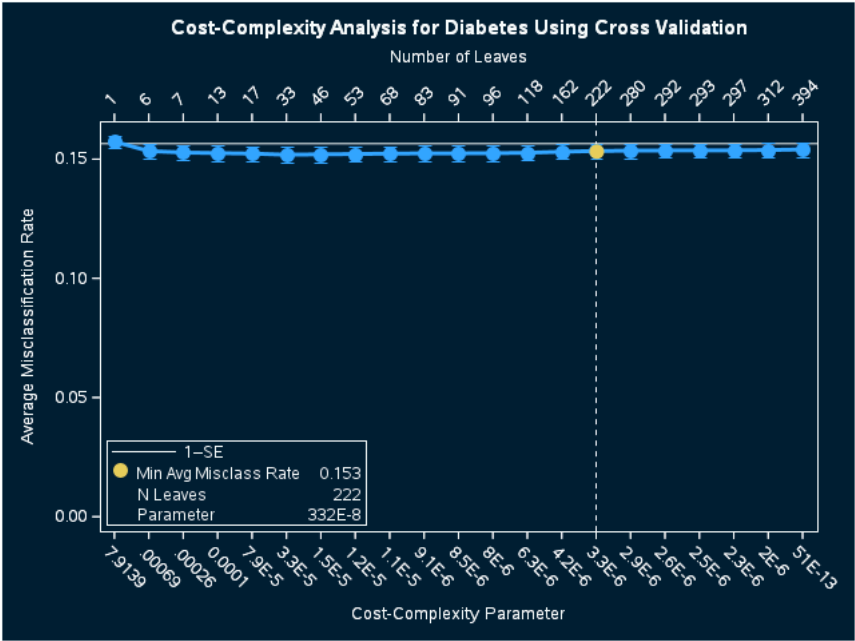
The HPSPLIT Procedure

Cost-Complexity Analysis for Diabetes Using Cross Validation

Diabetes Health Indicators

The HPSPLIT Procedure

Classification Tree for Diabetes

Subtree Starting at Node=0

1 Diabetes=0   2 Diabetes=1   3 Diabetes=2
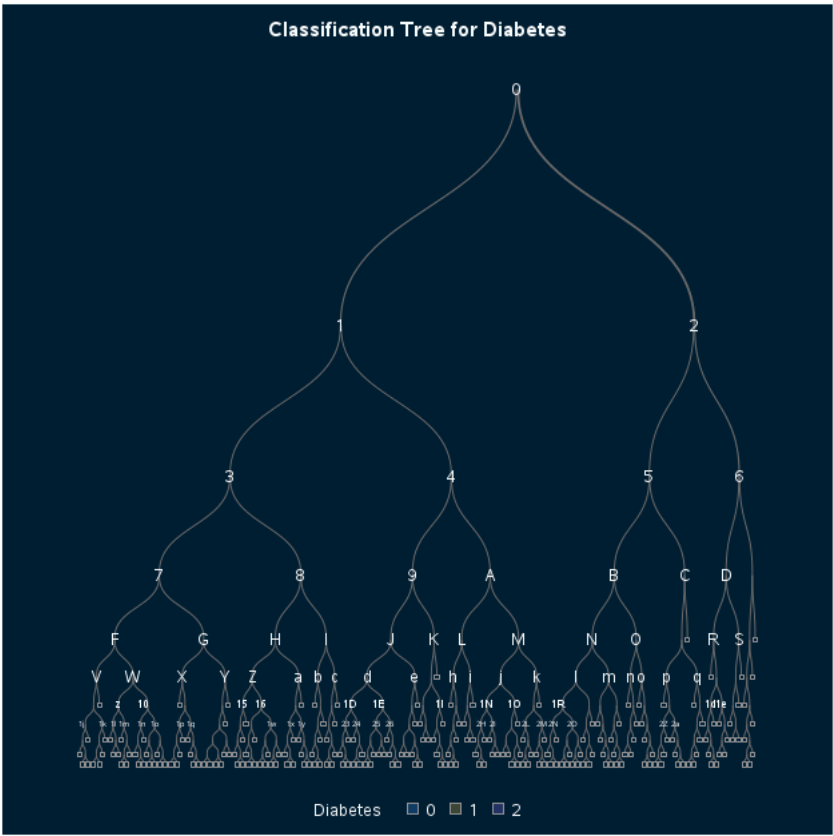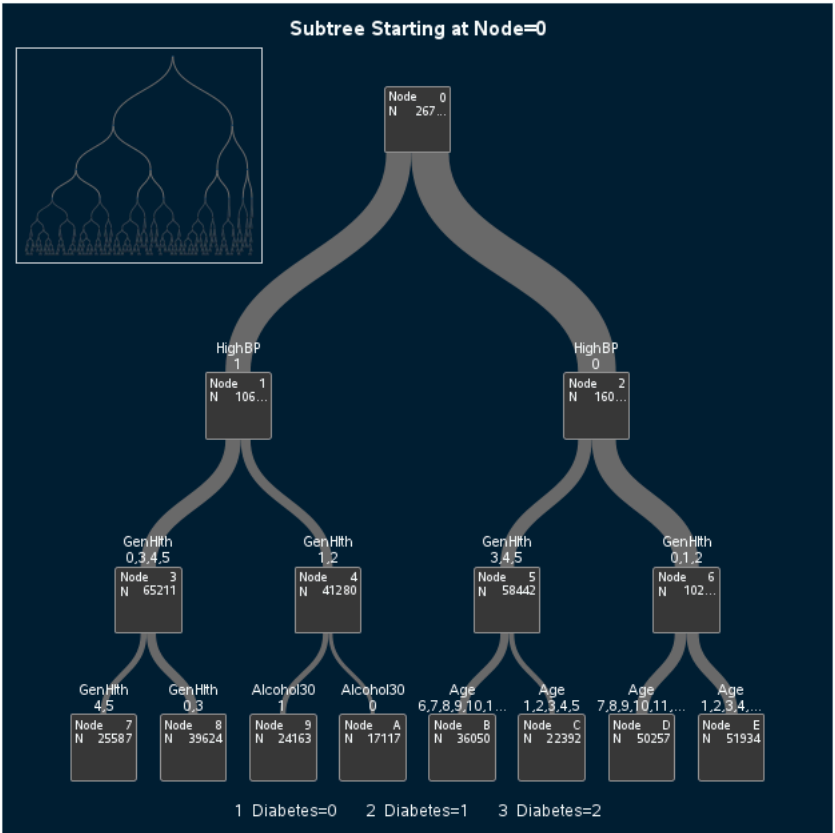
Diabetes Health Indicators

The HPSPLIT Procedure

Model-Based Confusion Matrix				
	Predicted			Error
Actual	0	1	2	Rate
0	220166	1	4924	0.0219
1	5514	24	612	0.9961
2	28852	2	7029	0.8041

Model-Based Fit Statistics for Selected Tree					
N Leaves	ASE	Mis-class	Entropy	Gini	RSS
219	0.0750	0.1494	0.5889	0.2249	60069.6

Variable Importance			
	Training		
Variable	Relative	Importance	Count
HighBP	1.0000	70.2226	1
GenHlth	0.8341	58.5752	8
Age	0.4699	32.9942	35
BMICat	0.4174	29.3085	13
HighChol	0.3742	26.2754	16
Alcohol30	0.2834	19.8999	7
CholCheck	0.2294	16.1108	10
Race	0.1985	13.9357	19
PhysHlth	0.1817	12.7589	35
MentHlth	0.1418	9.9541	25
Income	0.1213	8.5198	21
HeartDiseaseorAttack	0.1170	8.2183	6
HvyAlcoholConsump	0.0986	6.9256	3
DiffWalk	0.0834	5.8556	3
Education	0.0660	4.6369	5

Variable Importance			
	Training		
Variable	Relative	Importance	Count
PhysActivity	0.0601	4.2202	3
Stroke	0.0595	4.1759	2
Sex	0.0558	3.9169	3
Veggies	0.0361	2.5354	1
Smoker	0.0288	2.0207	1
CurrESmoke	0.0273	1.9169	1