

## SUPPORT Applications for Classification Trees

Sang Bock Lee<sup>1)</sup> · Sun Young Park<sup>2)</sup>

### Abstract

Classification tree algorithms including as CART by Brieman et al.(1984) in some aspects, recursively partition the data space with the aim of making the distribution of the class variable as pure as within each partition and consist of several steps. SUPPORT(smoothed and unsmoothed piecewise-polynomial regression trees) method of Chaudhuri et al(1994), a weighted averaging technique is used to combine piecewise polynomial fits into a smooth one. We focus on applying SUPPORT to a binary class variable. Logistic model is considered in the caculation techniques and the results are shown good classification rates compared with other methods as CART, QUEST, and CHAID.

**Keywords** : CART, CHAID, Classification tree, QUEST, SUPPORT

### 1. 서론

의사결정나무는 의사결정규칙을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류하거나 예측을 수행하는 분석방법이다. 의사결정나무는 분류 또는 예측을 목적으로 하는 어떤 경우에도 사용될 수 있으나 분석의 정확도 보다는 분석 과정의 설명이 필요한 경우에 더 많이 사용된다.

의사결정나무의 장점은 첫째, 나무 구조에 의해서 모형이 표현되기 때문에 모형을 사용자가 쉽게 이해 할 수 있다. 또한, 새로운 개체에 대한 분류 또는 예측을 하기 위해서 뿌리노드(root node)로부터 끝노드까지를 단순히 따라가면 되기 때문에, 새로운 자료에 모형을 적합시키기가 매우 쉽다. 둘째, 비모수적인 방법으로서 어떤 특정한 함수 형태를 요구하지도 않고, 모수적인 모형처럼 선형성, 정규성 또는 등분산성등의 가정을 필요로 하지 않는다. 따라서 독립변수의 변환없이 그 자체를 사용할 수 있다. 셋째, 독립변수 선택을 미리 하지 않아도 된다. 따라서 전진, 후진, 단계적 변수 선택방

---

1) 제1저자 : Professor, Dept. of Information & Statistics, Catholic University of Daegu, 712-702.

E-mail: sangbock@cu.ac.kr

2) Lecturer, Dept. of Information & Statistics, Catholic University of Daegu, 712-702.

법 등이 필요하지 않으며 모든 가능한 독립변수를 사용할 수 있다. 넷째, 결측치나 이상치를 가진 자료를 잘 처리 할 수 있다. 의사결정나무는 주분류변수가 결측치를 가진 어떤 자료를 분류할 때, CHAID(chi-squared automatic interaction detection)에서 처럼 결측치를 또 다른 범주로 생각하던지 또는 CART(classification and regression trees)처럼 다른 분류변수를 사용하여 결측치를 처리하고 있다. 이상치에 대해서는 의사결정나무의 특성상 하나의 노드로 분류되어진다. 그러나 이러한 장점에도 불구하고 몇 가지 단점이 있는데, 첫째, 불안정성이다. 조그마한 자료의 변화에도 나무구조는 많이 달라 질 수 있다. 이것은 나무구조의 국소최적설계 때문이다. 그러나 일반적으로 상위단계의 구조모형은 크게 다르지 않고, 하위단계에서 조금씩 달라질 뿐이며 전체 나무의 정확도에는 큰 차이가 없다. 둘째, 분류값의 주변에 대한 엄격한 분류이다. 실제로 의사결정나무에서는 연속형 변수를 비연속적인 값으로 취급하기 때문에 분리의 경계점 근방에서는 예측오류가 클 가능성이 있다. 최근에 이러한 단점을 극복하기 위하여, 앞의 장점을 해치지 않고 모수적 모형이나 신경망 등을 의사결정나무와 결합하는 방법들이 연구되고 있다. 셋째, 회귀모형에서는 회귀계수나 오즈비를 이용하여 결과에 대한 유용한 해석을 얻을 수 있다. 즉, 선형모형에서 주효과는 다른 예측변수와 관련시키지 않고서도 각 변수의 영향력을 해석할 수 있다는 장점을 가지고 있는데 의사결정나무에서는 선형 또는 주효과 모형에서와 같은 결과를 얻을 수 없다는 한계점이 있다.

의사결정나무분석을 수행하기 위한 다양한 분리기준, 정지규칙, 가지치기 방법들이 제안되어 있으며, 이들을 어떻게 결합하느냐에 따라서 서로 다른 의사결정나무 형성 방법이 만들어진다. 또한 정확하고 빠르게 의사결정나무를 형성하기 위해서 다양한 알고리즘이 제안되어 있고, 보다 개선된 알고리즘들이 계속 연구되어 발표되고 있다. 그 가운데 잘 알려진 알고리즘으로는 CHAID, CART, C4.5(Quinlan(1993)), QUEST(quick, unbiased, efficient, statistical tree)등이 있다. 의사결정나무는 목표변수의 성질에 따라 크게 두 가지로 나뉜다. 목표변수가 범주형 변수일 때는 분류의사결정나무라 하며, 연속형 변수일 때는 회귀의사결정나무라 한다. 분류의사결정나무에서 몇 가지 제안된 알고리즘을 보면 분류나무에서의 관심노드 분류기준법은 이영섭(2001)에 의해 제안되었고, CHITES(chi-square test and exhaustive search) 방법과 F&CHITES(F & chi-square test and exhaustive search) 방법은 Lee와 Song(2002)에 의해 제안되었다. 회귀의사결정나무에서 제안된 알고리즘을 보면 일반화된 회귀나무는 Chaudhuri 등(1995)에 의해 연구되었다. 그리고 회귀의사결정나무에서의 관심노드 찾는 분류는 이영섭(2003)에 의해 연구되었다.

이 논문의 목적은 목표변수가 이항일 경우에 사용하는 다중 로지스틱 회귀모형에 자료를 적합시켜 추정한 다음 나무 구조로 된 회귀의 새로운 방법인 SUPPORT(smoothed and unsmoothed piecewise-polynomial regression trees, Chaudhuri et al.(1994))를 이용하여 분류하는 것이다. 그리고 함수 추정이 어떻게 평활로 되는지 알아본다. 이 기술은 미지 회귀 함수에 대해 여러 도함수의 추정을 제공한다. 제안된 방법은 세 가지 원리를 가지는데, 하나는 각 부분집합은 다항 모형 적합을 이용하고, 그 다음 회귀공간의 이웃하는 부분 집합을 생성하기 위해 새로운 순환하는 분할 알고리즘을 이용하고, 마지막으로 한 개의 평활한 것으로 불연속 다항 회귀 추정을 결합하기 위해 가중평균을 이용한다. 여기서 얻은 의사결정나무와 다른 알고리즘들 즉, CART, QUEST, CHAID 각각의 의사결정나무를 비교하고 또한 이들에 대해 오분류

을 비교해 보고자 한다.

## 2. SUPPORT 알고리즘을 이용한 분류분석

회귀모형의 SUPPORT 추정은 자료 공간의 각 구간에서 관측치들에 대해 다항회귀 적합에 의해 얻어진다. 분할은 나무 구조로 된 방법으로서 순환적으로 실행된다. 평할은 다항 구간들의 가중 평균에 의해 이루어지며, 세 단계를 가진다. 첫 번째 단계에서, 회귀 공간은 각 구간에 있는 자료가 고정된 차수의 다항식에 의해 적당하게 적합될 때까지 순환적으로 분할된다. 분할은 잔차 분포의 분석과 PMSE(prediction mean square error)의 교차타당성 추정에 의해 얻어진다. 두 번째 단계는 각 분할의 이웃안의 자료는 다항에 의해 적합된다. 회귀함수의 마지막 추정은 다항구간들의 평균에 의해 얻어지는데, 평할가중함수의 각각은 연관된 분할 밖에서는 0으로 된다.

회귀분석에서, 우리는 전형적으로 목표변수  $Y$ 와  $k$ 개 독립변수  $X=(X_1, \dots, X_k)$ 에서  $n$ 개 관측치들을 가진다. 목표변수는  $Y=g(X)+\varepsilon$  을 통해  $X$ 에 의존한다고 가정된다. 이때  $\varepsilon$ 은 모형에서 오차이며, 이것은 어떤  $X$ 가 주어질 때 평균 0을 가지는 확률변수라고 가정된다. 회귀분석의 목적은 어떤 손실함수를 최소화하는  $g(X)$ 의 추정치  $\hat{g}(X)$ 를 찾는 것이다. 때때로  $g$ 가 특정형태를 가지는 것을 가정하기에 적당하다.

$$g(X) = \beta_0 + \sum_{k=1}^K \beta_k h_k(X_k)$$

이때 함수  $h_1(X_1), \dots, h_k(X_k)$  값은 알려진다. 단순한  $h$ 로 일차함수를 들 수 있다. 그러나 계수  $\beta_k$ 는 미지수다. 이것은 모수적 회귀 문제의 예제로써, 위 식은 미지 모수의 고정된 수를 포함하고 표본크기  $n$ 에 대해 변하지 않는다. 비모수회귀분석은  $g$ 에서 더 약한 가정을 만들어서 모수 설명을 일반화한다. 회귀모형  $g$ 의 추정을 얻는 방법은, B-스플라인과 평할 스플라인, 커널 평할기, 그리고 국소적 가중 회귀를 포함한다. 다른 방법들은 평할 함수의 합을 가지고 회귀모형  $g$ 근사에 접근한다. 예를들면 사영추적회귀(Freidman 등(1981)), 일반화된 부가적 모형(Buja 등(1989)), Hastie와 Tibshiani(1990), Stone(1985)), 그리고 PIMPLE(Breiman (1991))등이 있다.

회귀모형  $g$ 의 추정은 자료와 회귀공간의 순환적인 분할에 의해 만들어진다.  $g$ 의 불연속 상수 추정을 하는 방법에는 AID(Sonquist (1970)와 CART가 있다. 그리고 이들은 Breiman과 Meisel(1976), 그리고 Friedman(1979)의 불연속 선형추정방법을 들 수 있다. 이것은 다른 방법이 가지지 않는 두 개의 큰 이점을 가지는데 하나는 순환하는 분할에 의해 만들어진 의사결정나무는 회귀변수들에 관해 유용한 정보를 제공할 수 있다는 것이고, 다른 하나는 각 노드에서 추정된 표면들은 단순하며 상수 또는 선형회귀와 같은 함수형을 가진다는 것이다. MARS(Friedman (1991))는 순환하는 분할을 가지고 스플라인 적합의 결합에 의해 연속적인 함수 추정을 만든다. 이 연속 함수가 만들어짐으로 생긴 손실은 모형 해석의 어려움이 증가된다는 것이다. 또 다른 어려움은, 그들이 복잡하기 때문에 MARS 추정의 통계적 성질을 분석하기에 매우 어렵다는 것이다.

SUPPORT 방법은 세 개 원리를 가지는데, 첫째는 각 부분집합을 적합하는데 다항 모형을 이용한다는 것이고, 둘째는 회귀 공간의 이웃하는 부분집합을 생성하기 위해 새로운 순환하는 분할 알고리즘(Huang (1989))을 이용하는 것이다. 마지막으로는 한 개의 평할한 것으로 불연속 다항 회귀 추정을 결합하기 위해 가중평균을 이용한다는 것이다(O'Sullivan (1991)).

CART와 SUPPORT의 특성을 <표 1>에서 요약하였다. 멈춤규칙과 목표변수가 연속이라는 점에서 CART와 차이가 있다.

<표 1> SUPPORT 와 CART 알고리즘 비교

CART	SUPPORT
각 마디에 상수 적용 표본크기가 증가함에 따라 단계수가 한계없이 증가하는 나무 생산	선형(또는 다항) 회귀 모형 적합 CART보다 더 짧은 나무 생성
지나치게 큰 나무의 역행하는 '가지치기'에 의한 회귀나무	나무크기를 결정하기 위해 교차타당성 다단계 앞으로 나아가는 멈춤규칙
잔차제곱합(RSS)에서 감소율을 기본으로 한 분리 선택	잔차분포의 분석에 의해 분리 선택

### 3. 피마 인디언의 당뇨 자료

예제는 Chaudhuri 등(1994)이 사용한 사례와 같은 변수구조인 이진 목적변수와 연속형 독립변수를 가진 피마인디언의 당뇨자료를 이용하였다. 여기서 E-Miner 경우와 Answer Tree에서 같은 CART 알고리즘의 표기 방법이 알고리즘을 사용하는 회사인 SAS사와 SPSS사에 따라 각각 CART와 C&RT로 다르다. 회사의 사용 표기대로 각각을 다르게 표기하였다.

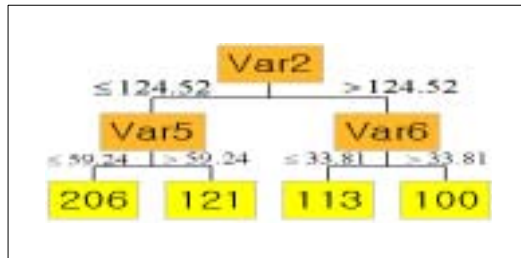
캘리포니아 주립대학 UC Irvine에 있는 자료 장소 (<http://www.ics.uci.edu/~mlearn/>)에서 피마 인디언 족의 자료를 가지고 비교, 분석하였다. 이 자료는 21세 이상의 피마족 여성 768명을 대상으로 그들의 당뇨병 여부를 목표변수로 하고, 당뇨병에 영향을 미칠 것이라고 생각되는 8개의 독립변수를 조사하였다. 768명 중 268명이 목표변수 값이 1인 경우인 당뇨병 환자이며 나머지는 목표변수 값이 0인 경우인 비당뇨병 환자이다. 자료 768개 중에서 학습표본 600개 그리고 테스트 표본 168개로 하였다. MINDAT값은 20, 10중 교차 타당성 추정의 PMSE는  $f$ 와  $\eta$ 를 0.1에서 0.5까지 0.1씩 증가시켜 얻어졌다. 그러나 여기서는 큰 차이가 나지 않았다. 그래서  $(f, \eta) = (0.2, 0.2)$ 로 하였다.

<표 2>은 피마 인디언 자료를 이용하여 다중 로지스틱 회귀모형에 적합시켜 얻은  $\beta$ 값과 표준오차 값이다. <그림 1>는 위 자료를 다중 로지스틱 회귀모형에 적합시킨 후 SUPPORT 알고리즘을 이용하여 얻은 의사결정나무다. 이때  $f, \eta$ 는 각각 0.2로

주었다. MINDAT는 20을 주었다. 가중치는  $\tau$ 를 0.25로 주었다. 각 끝 노드에는 표본 크기가 쓰여져 있다. 첫 번째 분리에서 선택된 변수는 V2인 혈당량이다. 그리고 두 번째 분리에서 선택된 변수는 V5인 인슐린 수치와 V6인 비만도이다. 168개의 테스트 표본 경우에 대해 추정된 PMSE는 0.183이다. SUPPORT로 계산된 확률들로 나타낼 때, 오분류율은 분류기준값을 목표변수가 2개의 범주를 가지므로 0.5로 주었을 때, 20%를 얻었다. 이러한 기준이 상식적이기는 하나 절대적인 기준이 될 수는 없다. 그래서 분류기준값을 0.25인 경우와 0.75인 경우에 대해서도 오분류율을 구해보았다. 분류기준값이 0.25인 경우는 오분류율이 27%, 분류기준값이 0.75인 경우는 오분류율이 26%였다.

<표 2> 피마 인디언 당뇨 자료

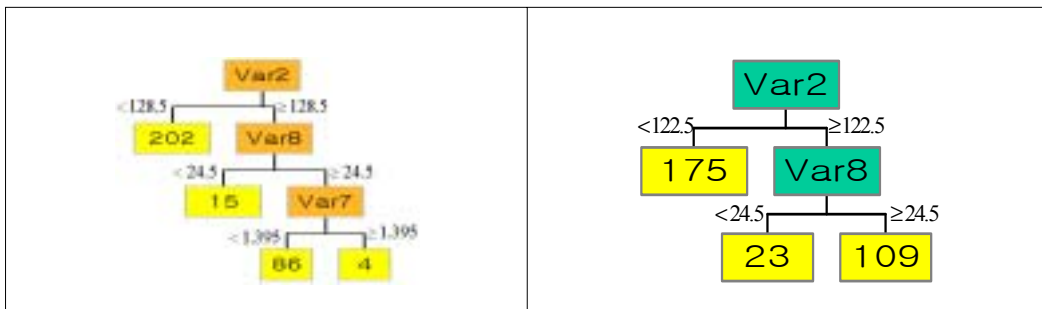
변수	추정계수	표준오차
const: 상수	-8.4042	0.7166
Var1(V1):number of times pregnant	0.1232	0.0321
Var2(V2):plasma glucose concentration a 2 hours in an oral glucose tolerance test	0.0352	0.0037
Var3(V3):diastolic blood pressure (mm Hg)	-0.0133	0.0052
Var4(V4):Triceps skin fold thickness (mm)	-0.0006	0.0069
Var5(V5):2-Hour serum insulin (mu U/ml)	-0.0012	0.0009
Var6(V6):Body mass index (Weght in kg/(heght in m)^2)	-0.0897	0.0151
Var7(V7):Diabets pedigree function	0.9454	0.2992
Var8(V8):Age (years)	0.0149	0.0093



<그림 1> 피마 인디언 자료에서 SUPPORT 나무

<그림 1>은 SAS의 E-miner를 이용하여 CART알고리즘 중 분리기준을 킨 후 SUPPORT 알고리즘을 이용하여 얻은 의사결정나무다. 이때  $f, \eta$ 는 각각 0.2로 주었다. MINDAT는 20을 주었다. 이때 가중치는  $\tau$ 를 0.25로 주었다. 각 끝 노드에는 표본 크기가 쓰여져 있다. 첫 번째 분리에서 선택된 변수는 V2인 혈당량이다. 그리고 두 번째 분리에서 선택된 변수는 V5인 인슐린 수치와 V6인 비만도이다. 168개의 테스트 표본 경우에 대해 추정된 PMSE는 0.183이다. SUPPORT로 계산된 확률들로

나타낼 때, 오분류율은 분류기준값을 목표변수가 2개의 범주를 가지므로 0.5로 주었을 때, 20%를 얻었다. 이러한 기준이 상식적이기는 하나 절대적인 기준이 될 수는 없다. 그래서 분류기준값을 0.25인 경우와 0.75인 경우에 대해서도 오분류율을 구해보았다. 분류기준값이 0.25인 경우는 오분류율이 27%, 분류기준값이 0.75인 경우는 오분류율이 26%였다. 지니 지수를 사용한 경우이다. 여기서 첫 번째 분리 변수는 V2인 혈당량이 선택되었고, 두 번째 분리 변수는 V8인 나이가 선택되었다. 오분류율은 24.3% 였다.

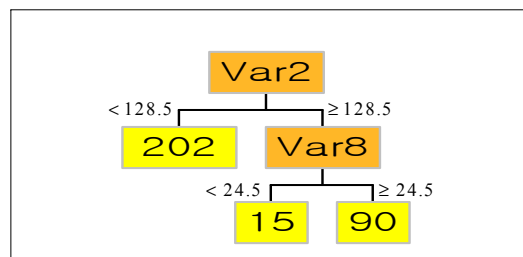


<그림 2> 피마 인디언 자료에서 CART 나무 1 <그림 3> 피마 인디언 자료에서 CART 나무 2

<그림 2>는 SAS의 E-miner를 이용한 CART 알고리즘 중 분리기준을 엔트로피 지수를 사용한 경우이다. 여기서 첫 번째 분리 변수는 V2가 선택되었고, 두 번째 분리 변수는 V8가 선택되었다. 오분류율은 26.1% 였다.

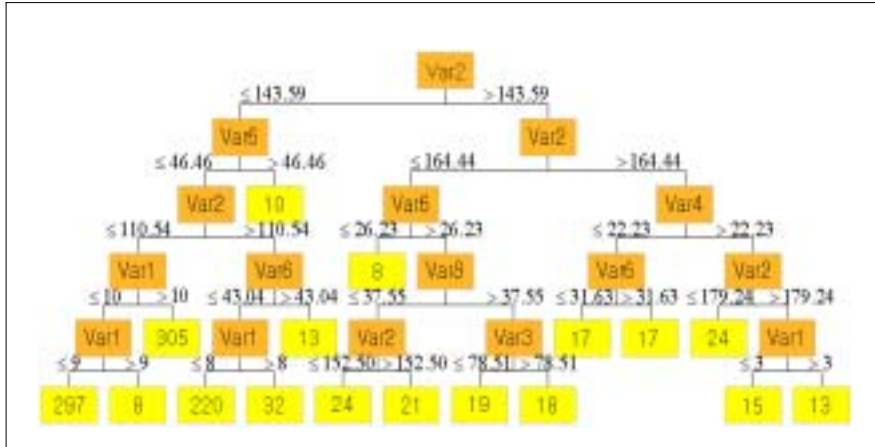
<그림 3>는 SAS의 E-miner를 이용한 CHAID 알고리즘을 사용한 경우이다. 분리 기준은 카이제곱 검정으로 하였고 유의수준은 0.05로 지정하였다. 여기서 첫 번째 분리 변수는 V2가 선택되었고, 두 번째 분리 변수는 V8이 선택되었다. 오분류율은 25.2% 였다.

<그림 4>은 SPSS의 Answer Tree에서 QUEST 알고리즘을 이용한 결과이다. 첫 번째 분리 변수는 V2가 선택되었고, 두 번째 분리 변수는 V6와 V2가 선택되었다. 오분류율은 23.6% 였다. <그림 5>과 <그림 6>는 SPSS의 Answer Tree에서 C&RT 알고리즘을 이용한 결과이다. 나무의 크기가 크기 때문에 첫 번째 분리 후 왼쪽 노드 이하 나무는 <그림 6>, 오른쪽 노드 이하 나무는 <그림 7>에 나타내었다.

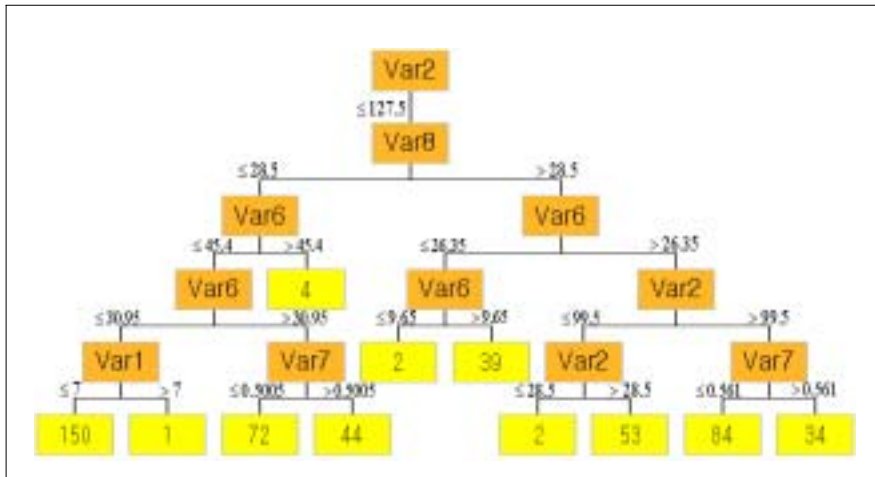


<그림 4> 피마 인디언 자료에서 CHAID나무

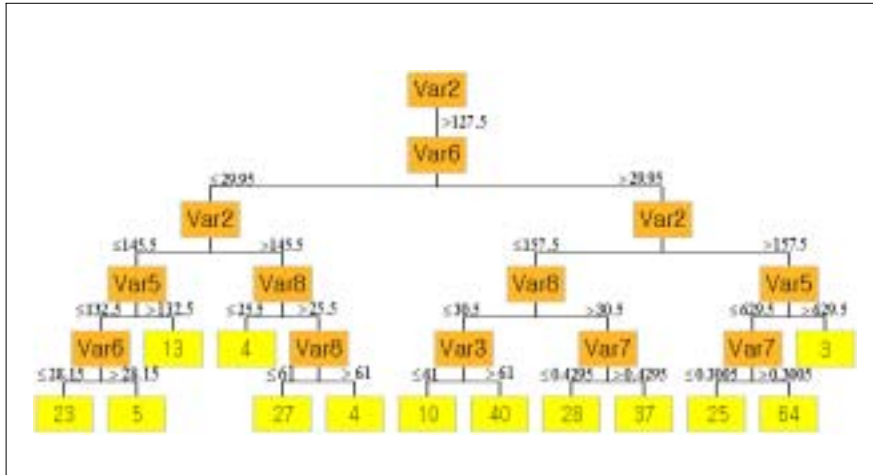
여기서 첫 번째 분리 변수는 V2 , 그리고 두 번째 분리 변수는 V6와 V8이 선택되었다. 오분류율은 16.5% 였다.



<그림 5> 피마 인디언 자료에서 QUEST 나무



첫 번째 분리 후 왼쪽 노드의 분리  
<그림 6> 피마 인디언 자료에서 C&RT 나무



첫 번째 분리 후 오른쪽 노드의 분리  
 <그림 7> 피마 인디언 자료에서 C&RT 나무(계속)

모든 알고리즘의 오분류율은 비슷하다고 볼 수 있는데 SPSS의 Answer Tree에서 C&RT 경우의 오분류율은 16.5%로 가장 작게 나타났다. 그러나 나무의 크기가 크다는 단점이 있다. CART 알고리즘은 이상치의 영향을 받아 이상치를 하나의 또 다른 그룹으로 분리하는 성질이 있기 때문에 첫 번째 변수 선택 후 두 번째에서는 모두 V8이 선택되었다. 변수 V8은 70대와 80대에서 이상치가 있는 자료이다. 그리고 SPSS의 Answer Tree에서는 CHEST와 C&RT 알고리즘에 대해 옵션 지정이 세분화 되지 않고 단순하기 때문에 형태가 큰 나무를 형성하는 단점이 있다.

#### 4. 결론

본 논문에서는 나무구조로 된 회귀방법인 SUPPORT에 분류분석을 연구하였다. 그리고 여기서 제안된 방법과 다른 알고리즘들의 의사결정나무와 오분류율을 비교해 보았다. SUPPORT의 경우 SAS E-miner를 이용한 의사결정나무와는 크기가 비슷하였으나 SPSS의 Answer-Tree를 이용한 경우보다는 훨씬 작은 나무구조를 형성하였다. 오분류율은 SUPPORT에서 알고리즘의 경우 분류기준값을 0.5로 주었을 때 SPSS의 Answer-Tree에서 C&RT 알고리즘을 제외하고는 다른 경우에 비해 조금 작게 나타났다. 그러나 CART 알고리즘의 경우 특별히 이상치가 있을 때 다른 하나의 노드로 분리하는 경향을 가지기 때문에 피마 인디언 자료에서 두 번째 분리변수로 나이변수 V8이 선택되었다. 그리고 SPSS의 Answer Tree는 C&RT에서 오분류율이 가장 낮게 나타나지만 옵션의 선택이 세분화 되어 있지 않기 때문에 CART와 QUEST 알고리즘들 다에서 큰 나무를 형성하는 단점이 있다.

나무 구조로 된 회귀에 대해 새로이 제안 한 기술은 구간별 상수 적합 대신에 다항 적합을 이용하였기 때문에 노드 수가 작은 나무를 형성한다. 또한 잔차 분석을 기본



으로 하는 이 방법은 CART의 전체탐색법보다 계산 시간이 짧고 구간별 다항 추정의 평할은 단순하고 효율적으로 볼 수 있다.

### 참고문헌

1. Breiman, L. (1991), The  $\Pi$  method for estimating multivariate functions from noisy data (with discussion), *Technometrics* 33,125-160.
2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J.(1984), *Classification and Regression Trees*, Wadsworth, Belmont.
3. Breiman, L. and Meisel, W.S.(1976), General estimates of the intrinsic variability of data in nonlinear regression models, *Journal of the American Statistical Association* 71, 301-307.
4. Buja, A., Hastie, T. and Tibshirani, R. (1989), Linear smoothers and additive models (with discussion), *Annals of statistics* 17: 453-555.
5. Chaudhuri, P., Huang, M. C., Loh, W. Y. and Yao, R.(1994), Piecewise-polynomial regression trees, *Statistica Sinica*, Vol. 4, 143-167.
6. Chaudhuri, P., Lo, W. D., Loh, W. Y. and Yang, C. C. (1995), Generalized Regression Trees, *Statistica Sinica*, Vol. 5, 641-666.
7. Freidman, J. H.(1979), A tree-structured approach to nonparametric multiple regression, in Gasser, T. and Rosenblatt, M.(eds), *Smoothing Techniques for Curve Estimation*, Springer-Verlag, 5-22, Lecture Notes in Mathematics 757.
8. Freidman, J. H. (1991), Multivariate adaptive regression splines (with discussion), *Annals of Statistics* pp. 1-141.
9. Friedman, J. H. and Stuetzle, W. (1981), Projection pursuit regression, *Journal of the American Statistical Association* 76: 817-823.
10. Hastie. T. and Tibshiani. R. (1986), Generalized Additive Models(with discussion), *Statistical Science* 1: 297-310.
11. Hastie. T. and Tibshiani. R. (1990), *Generalized Additive Models*. Chapman and Hall, London.
12. Huang, M. C. (1989), Piecewise-linear Tree-structured Regression, Ph.D. thesis, University of wisconsin, Madison, Department of Statistics.
13. Lee, Y. M. and Song, M. S. (2002), A Study on Unbiased Methods in Constructing Classification Trees, *The Korean Communications* Vol. 9, 809-824.
14. O'Sullivan(1991), Discussion of 'Multivariate adaptive regression splines' by J.H. Friedman, *Annals of Statistics*, 19, 99-102.
15. Quinlan, J.(1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Mateo.
16. Sonquist, J. (1970), Multivariate model building: The validation of a search strategy, Technical report, Institute for Social Research,

University of Michigan, Ann Arbor.

17. Stone, C. J. (1985), Additive regression and other nonparametric models, *Annals of Statistics* 13: 689-705.
18. 박선영(2004), SUPPORT를 이용한 분류분석 연구, 대구가톨릭대학교 대학원, 박사학위논문.
19. 이영섭. (2001), New Splitting Criteria for Classification Trees, *The Korean Communications in Statistics*, Vol. 8. No. 3. 885-894.
20. 이영섭. (2003), 회귀의사결정나무에서의 관심노드 찾는 분류기준법, *응용통계연구* 제 16권 1호, 45-53.

[ 2004년 5월 접수, 2004년 8월 채택 ]