

# Decision-Tree를 이용한 터널 보조공법 의사결정 Decision-making of Tunnel Reinforcement Methods using Decision-Tree Technique

김종규\*      이준석\*\*      사공명\*\*\*      이용주\*\*\*  
J. G. Kim      J. S. Lee      Myung Sagong      Y. J. Lee

---

## ABSTRACT

Auxiliary reinforcement methods such as for-poling and multi-steel pipe grouting etc. are frequently used to avoid a possible collapse of tunnel face, where the ground condition is not favourable as well as geological formation is not sufficient. However, among those methods a suitable reinforcement method is exclusively decided by tunnel engineers' experience so far. Therefore, this study introduces a data-mining technique for the proper auxiliary reinforcement method during uncertainty of tunnelling work. The data-mining technique has widely been used in the field of social science, medical treatment, banking and agriculture etc. A decision-tree together with PDA technique based on tunnel construction data was developed to provide a reasonable reinforcement method in this study. Future study will be focused on the application of this technique in a variety of underground space developments.

---

## 1. 서론

터널의 보조공법은 터널굴착에 따른 원지반, 주변 구조물, 막장면 등의 안정을 위해 적용된다. 그러나 지반의 불확실성으로 인해 뚜렷한 보조공법을 제시하기가 어렵고, 정량적인 평가항목을 정하는데 어려움이 있어 터널현장 기술자의 경험을 바탕으로 보조공법을 결정하는 경우가 빈번히 발생하게 된다.

본 연구에서는 데이터 마이닝 (Data mining) 을 이용한 터널 보조공법의 선정기법과 선정과정의 자동화를 위한 PDA S/W를 개발하였다. 데이터 마이닝은 대용량의 자료로부터 사용목적에 부합한 유용한 정보를 선별하는 기법으로 데이터 사이에 숨겨진 패턴을 발견하고 규칙을 추론함으로써, 의사결정을 지원하는 기법이다. 특히, 지난 수십 년간 축적된 전자상거래, 주식거래, 은행거래, 신용카드결제 및 세금 등의 대용량 자료의 경우, 데이터의 양은 계속 증가하는 상황에서 의미있는 정보를 찾아내는 과정을 데이터마이닝이라 부를 수 있다. 과거의 경우 제한된 정보범주에서는 단지 의사 결정자의 직관에 근거하여 전문가의 지식을 지식베이스에 수동으로 입력하였기 때문에 왜곡과 오류를 낳기 쉬운 반면, 데이터마이닝을 적용하여 중요한 데이터 패턴을 찾아내고 보다 객관적이고 효율적인 의사결정을 수행할 수 있다. 따라서 본 연구에서는 국내 시공된 터널 설계데이터를 근거로 하여 터널 보조공법 선정분야에 데이터 마이닝 기법을 적용하였고, 터널 보조공법에 관련된 인자의 교호효과를 분석하여 터널 시공 중 현장 상황에 따른 보조공법을 제시할 수 있는 방안을 연구하였다.

---

\* 한국철도기술연구원 연구원, 비회원

\*\* 한국철도기술연구원 책임연구원, 정회원

\*\*\* 한국철도기술연구원 선임연구원, 정회원

## 2. 의사결정트리

의사결정트리분석은 대용량의 데이터 내에 존재하는 관계, 패턴, 규칙 등을 탐색하고 모형화하는 데이터마이닝 기법으로 신경망이나 판별분석 등에 의한 방법과는 달리 적용결과에 의해 규칙을 명확하게 나타내는 기법으로 트리구조로 표현되기 때문에 해석이 용이하며, 유용한 입력변수를 찾아내어 입력변수가 결합하여 목표변수에 어떻게 영향을 주는지 교호효과를 분석할 수 있는 알고리즘이다. 특히 의사결정트리의 가장 큰 장점은 이해하기 쉬운 규칙과 모형을 구축하는 것이다. 트리구조에 의해서 모형이 표현되기 때문에 모형을 사용자가 쉽게 이해할 수 있으며, 어떤 데이터가 결과를 설명하기 위해서 더 중요한지를 쉽게 파악할 수 있는 장점이 있다. 한편 연령과 같은 연속형 데이터를 처리하는데 있어 비연속적인 값으로 취급하기 때문에 예측오류가 발생할 수 있으며, 분석용 데이터에만 의존하기 때문에 분석용 자료의 크기가 너무 작거나 많은 가지로 구성된 모형에서 불안정한 규칙을 추론할 수 있다. 그러나 이러한 불안정한 규칙은 의사결정트리의 검증 및 타당성 분석을 통해 안정성 있는 모형을 구축할 수 있다. <그림 1>은 의사결정트리 분석의 수행과정을 나타내고 있다.

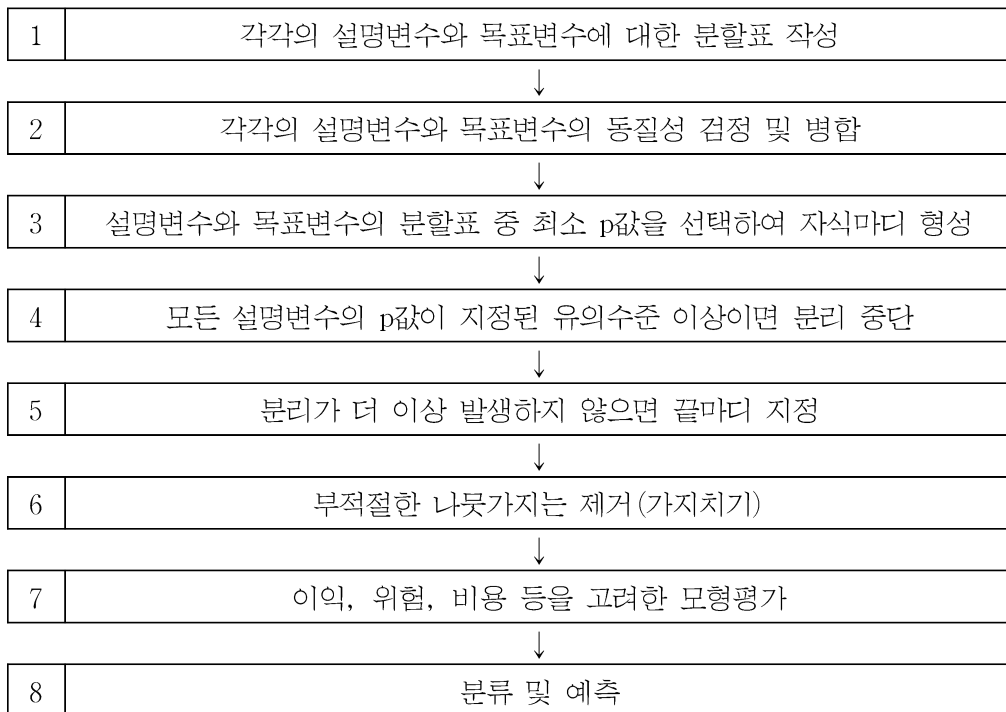


그림 1. 의사결정트리의 평가과정

<그림 2>는 의사결정트리구조로서 마디(node)로 구성되며, 부모마디(parent node)로부터 시작하여 하위마디로 자식마디(child node)를 계속적으로 분리, 형성해 나감으로써 완성된다. 부모마디와 반대로 트리의 가장 끝에 위치하여 가지가 분리되지 않는 마디를 끝마디(종단마디, leaf node)라고 하며, 부모마디부터 종단마디까지의 분리단계를 깊이(depth)라고 한다. <그림 2>에서 총 6개의 최종마디가 생성되었고 상위마디의 규칙을 가정-결론(IF-THEN) 방식으로 추론할 수 있다. 예를 들어, 최종마디 “신용카드유”의 경우 “IF 나이=35세 이상 AND 월소득=200만원 이상 AND 신용카드=유 THEN 신용상태=GOOD”와 같은 규칙이 추론된다.

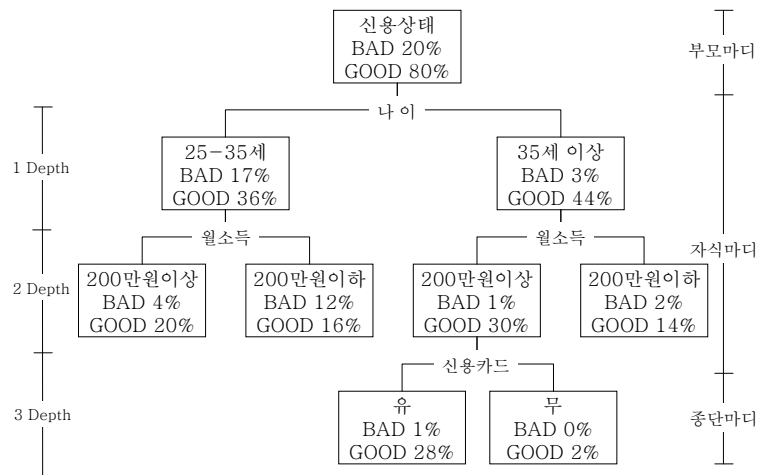


그림 2. 의사결정트리 (강현철 외, 2001)

지금까지 의사결정트리 분석을 위해 많은 알고리즘이 제시되었고, 대표적인 것들로 CHAID (Kass, 1980), CART (Breiman et al., 1984), C4.5 (Quinlan, 1993), QUEST (Loh and Shih, 1997) 등이 있다. 이들 중 CHAID (Chi-squared Automatic Interaction Detection)는 카이제곱-검정 (이산형 목표변수) 또는 F-검정 (연속형 목표변수)을 이용하여 분리와 병합을 반복하면서 Pearson 카이제곱 또는 우도비 (likelihood ratio) 카이제곱 통계량을 분리기준으로 사용한다. 카이제곱 통계량은 관측도수 ( $f_{ij}$ )로 이루어진  $r \times c$  분할표로부터 계산되며 분할표의 구조는 <표 1>과 같다.

표 1. 카이제곱 통계량의 분할표

목표변수 설명변수	범주1	범주2	...	범주 $c$	합계
범주 1	$f_{11}$	$f_{12}$	...	$f_{1c}$	$f_{1.}$
범주 2	$f_{21}$	2	...	$f_{2c}$	$f_{2.}$
...	...	...	...	...	...
범주 $r$	$f_{r1}$	$f_{r2}$	...	$f_{rc}$	$f_{r.}$
합계	$f_{.1}$	$f_{.2}$	...	$f_{.c}$	$f_{..}$

분할표로부터 Pearson의 카이제곱 통계량은 다음과 같다.

$$x^2 = \sum_{i,j} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \quad \text{-----} \quad (1)$$

분할표로부터 우도비 카이제곱 통계량은 다음과 같다.

$$x^2 = 2 \sum_{i,j} f_{ij} \times \log_e \left( \frac{f_{ij}}{e_{ij}} \right) \quad \text{-----} \quad (2)$$

두 통계량의 자유도 (degree of freedom)는  $(r-1)(c-1)$ 이며,  $e_{ij}$ 는 분포의 통일성 또는 독립성의 가설 하에서 계산된 기대도수 (expected frequency)를 말하며 다음과 같다.

$$e_{ij} = \frac{f_{i.} \times f_{.j}}{f_{..}} \quad (3)$$

카이제곱 통계량이 자유도에 비해서 매우 작다는 것은 설명변수의 각 범주에 따른 목표변수의 분포가 서로 동일하여 설명변수가 목표변수의 분류에 영향을 미치지 않는 것이다. 그리고 자유도에 대한 카이제곱 통계량이 크고 작음은 해당집단에서 그 조건이 틀렸다고 잘못 생각할 확률인 p값으로 표현되는데, 카이제곱 통계량이 자유도에 비해 작으면 p값은 증가하므로, p값이 가장 작은 설명변수에서 최적분리가 발생하는 것이다.

### 3. 터널 보조공법의 의사결정트리 분석

#### 3.1 터널 보조공법의 입력변수와 목표변수

터널단면의 계획, 굴착방법의 선정, 지보형식의 결정 등을 지배하는 주요 요소는 암반상태(암종, 풍화도, 절리, 파쇄대 등의 불연속면의 발달상태 등), 역학적 특성(압축강도, 인장강도, 전단강도 등), 변형특성(포아송비, 토압계수, 탄성계수 등) 및 지하수 상태 등이다. 이들 요소 중 터널의 굴착방법이나 지보형식을 선정하는데 중요한 사항은 암반상태와 역학적 특성이며, 특히 지보형식을 결정하기 위한 암반의 분류는 암질의 상태에 의존하며, 터널굴착 중 보조공법 적용시 암질의 상태는 중요한 항목이다(한국철도시설공단, 2004). 그리고 터널 기술자가 터널의 안정성을 쉽게 판단할 수 있는 조사항목이 선정되어야 하며 특히 시시각각 변하는 현장 상황에 유연하게 대처하기 위해 조사시간, 조사자의 이해력을 고려하여 복잡하지 않아야 한다. 또한 단순한 조사항목을 선정하기 위해 검증되지 않는 평가기준을 도입할 수는 없으므로 터널의 막장면 매핑자료를 기준으로 검증된 조사항목을 선정하여야 한다. <표 2>는 터널 보강계획을 수립하는데 있어서 보강목적에 적합한 공법을 선정하기위해 해당 지점의 지반조건과 시공환경을 고려한 조사항목을 나타내고 있다.

표 2. 보조공법의 입력변수와 목표변수

변수이름	입력/목표	변수값(코딩값)
RMR	입력	100-40(1), 40-20(2), 20-0(3)
토피고	입력	2D이상(1), 2D이하(2)
단층유무- 단층이격거리	입력	유-1m이하(1), 유-2D이하(2), 유-3m이상(3), 무(0)
인접구조물	입력	유(1), 무(0)
시공환경	입력	도심지(1), 산악(2)
갱구부	입력	해당(1), 해당없음(0)
보조공법	목표	Fore-poling, 강관다단 보강그라우팅, 대구경 강관다단 보강그라우팅, Fore-poling+강관다단 보강그라우팅

2000년 이후의 터널 설계자료를 바탕으로 이미 시공이 되어있거나 시공 중인 지하철, 고속국도 및 지방도, 고속철도의 터널 중 보조공법을 적용한 구간의 설계 데이터를 수집하여 <표 2>의 변수에 따라 분류하였다. 수집된 자료는 총 151개로 구성되었으며, 최종 목표변수는 Fore-poling 및 강관다단 보강그라우팅, 대구경 강관다단 보강그라우팅, 조합공법의 보조공법으로 <표 3>은 각각의 구성을 나타내고 있다.

표 3. 보조공법별 데이터 구성

구분	Fore-poling	강관다단 보강그라우팅	대구경 강관다단 보강그라우팅	Fore-poling + 강관다단 보강그라우팅	계
도심지터널	23	26	24	10	83
산악터널	38	26	4	0	68
계	61	52	28	10	151

### 3.2 보조공법의 의사결정트리

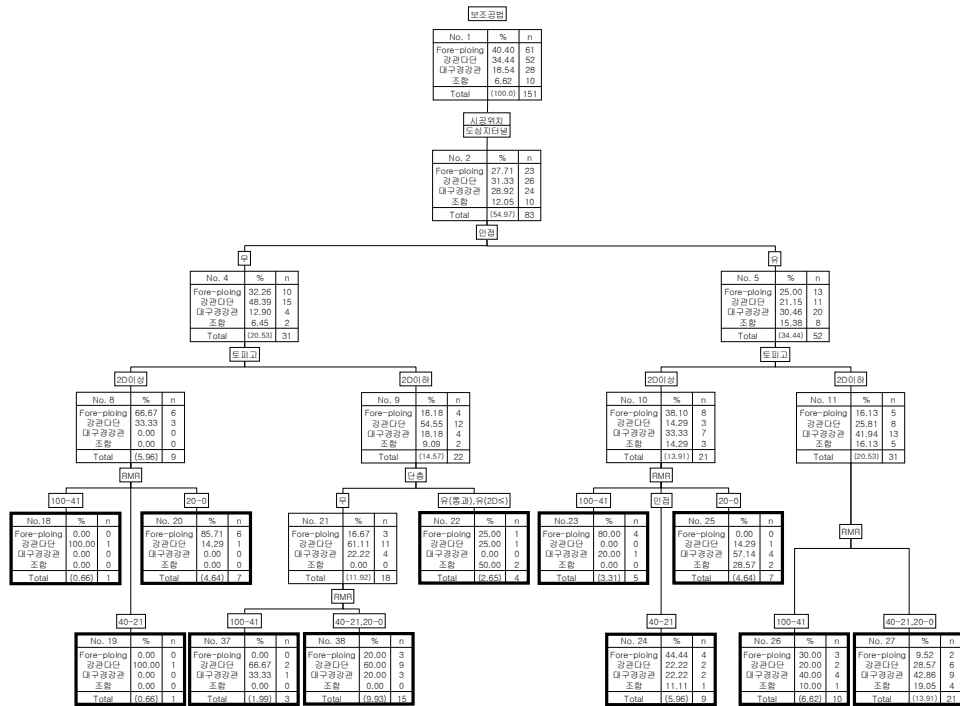
〈표 2〉에서 각 입력변수와 목표변수는 명목형변수, 즉 성별, 상품의 종류, 지역 등과 같이 순서를 가지지 않는 변수들로 구성되었기 때문에 변수별 코딩형식을 추가하여 의사결정트리의 모형을 구축하였다. 보조공법에 미치는 입력변수의 영향을 분석하기 위해 시공위치, 갱구부, RMR, 토피고, 단층, 인접구조물의 6개 입력변수를 이용하였고, 유의수준은  $p < 0.15$  로 하였다. 세부적인 알고리즘은 CHAID법을 이용하여 의사결정트리의 모형을 구축한 후 위험도표의 분석을 통해 안정성 있는 모형인지의 적합도를 검증하였다.

〈그림 3〉은 보조공법의 의사결정트리 결과로서 부모마디부터 최종마디까지 총 5단계로 분류되었으며, 부모마디를 포함하여 42개의 마디로 구성되었다. 굵은 실선으로 표시된 24개의 최종마디 중 도심지터널에 대한 최종마디는 11개, 산악터널에 대한 최종마디는 13개로 구분된다. 도심지터널과 산악터널의 첫 번째 분류항목은 시공변수로 나타나고 있으며, 대구경 강관다단 보강그라우팅의 경우 산악터널이 16% (4/28)인 반면 도심지터널은 86% (24/28)가 적용되었다. 또한 토피고와 인접구조물에 대한 전체 입력데이터의 구성은 도심지터널에서 토피고가 20m이하인 경우 64% (53/83), 인접구조물이 있는 경우 63% (52/83)인 반면에 산악터널은 34% (23/68), 24% (16/68)로 도심지터널의 특수한 시공환경이 전체 데이터와 의사결정트리에 반영된 것으로 판단된다.

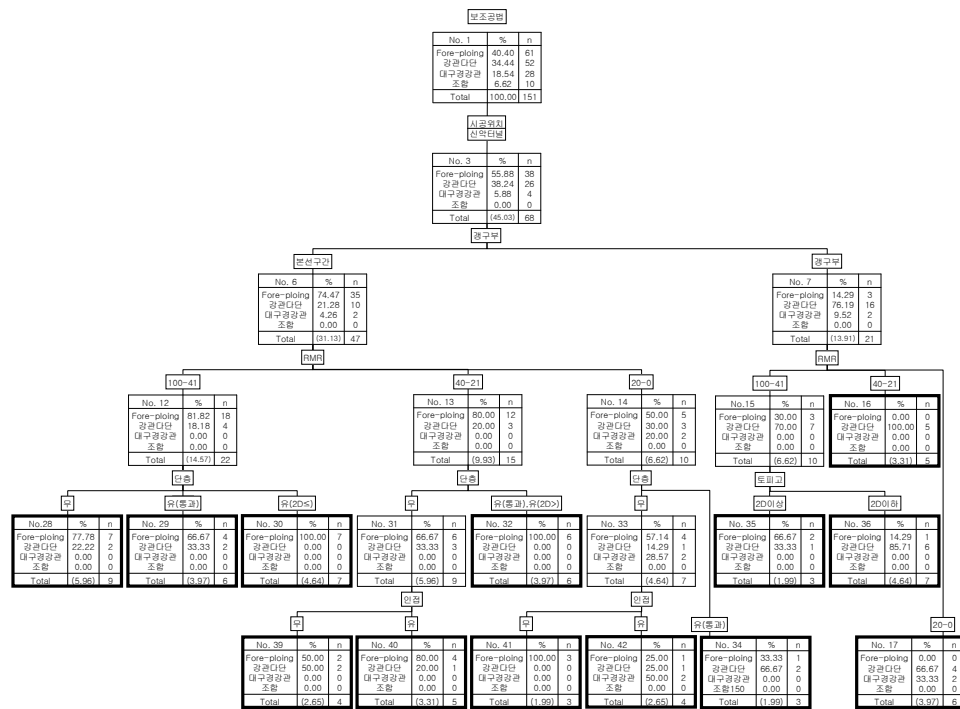
도심지터널의 의사결정트리는 시공환경, 인접구조물의 유, 무와 토피고에 따라 분류되었으며, 특히 인접구조물이 없고 토피고가 2D이상인 경우 주로 Fore-poling공법이 적용되었다. 반면 인접구조물이 있고 토피고가 2D이하인 경우 주로 강관다단 보강 그라우팅공법과 대구경 강관다단 보강 그라우팅공법이 적용되었다.

산악터널은 도심지터널과 달리 갱구부에 따라 의사결정트리가 분류되었으며, 이는 도심지터널과 다른 시공조건이 반영된 것이라 판단된다. 특히 갱구부에 해당하는 산악터널은 RMR이 유리한 경우 토피고에 따라 보조공법이 결정되었다. 그리고 RMR이 불리한 경우, 즉 RMR값이 0-20일 때 단층의 유, 무에 따라 결정되며, 단층이 없는 구간은 인접구조물의 유, 무에 따라 보조공법이 결정되었다. 또한 RMR이 유리한 60-100구간에서는 단층과 인접구조물의 영향으로 보조공법이 적용되었다.

〈그림 3〉의 의사결정트리에서 24개의 최종마디는 최종적으로 “IF-THEN” 형식의 규칙으로 표현되며 하나의 규칙은 부모마디에서 부모마디까지의 각 경로에 해당하며 〈표 4〉와 같다. 추론규칙을 정하기 애매한 마디에 대해서는 각 보조공법에 대한 해당 부모마디의 분포가 5% 이하인 경우 자료의 오차범위를 고려하여 추론규칙에서 제외하였다. 그리고 부모마디에서 보조공법의 분포가 동일하여 우위가 결정되지 않을 경우 해당하는 보조공법을 모두 적용시키도록 추론규칙을 정하였으며, 단 해당마디의 보조공법이 전체보조공법에 대하여 10% 이하인 경우 추론규칙으로 정하지 않고 현장조건에 따라 취사선택할 수 있도록 하였다.



(a) 도심지터널



(b) 산악터널

그림 3. 보조공법의 의사결정트리

표 4. 보조공법의 추론규칙

Node No.	IF	THEN
18	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이상" AND RMR="100-41"	미적용
19	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이상" AND RMR="40-21"	미적용
20	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이상" AND RMR="20-0"	Fore-poling
22	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이하" AND 단층="유(통과, 2D이하)"	Fore-poling + 강관다단그라우팅
37	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이하" AND 단층="무" AND RMR="100-41"	강관다단그라우팅
38	시공조건="도심지터널" AND 인접구조물="무" AND 토피고="2D이하" AND 단층="무" AND RMR="40-21, 20-0"	강관다단그라우팅
23	시공조건="도심지터널" AND 인접구조물="유" AND 토피고="2D이상" AND RMR="100-41"	Fore-poling
24	시공조건="도심지터널" AND 인접구조물="유" AND 토피고="2D이상" AND RMR="40-21"	Fore-poling 또는 Fore-poling +강관다단그라우팅
25	시공조건="도심지터널" AND 인접구조물="유" AND 토피고="2D이상" AND RMR="20-0"	대구경 강관다단 그라우팅 또는 Fore-poling +강관다단그라우팅
26	시공조건="도심지터널" AND 인접구조물="유" AND 토피고="2D이하" AND RMR="100-41"	대구경 강관다단 그라우팅, 필요시 Fore-poling
27	시공조건="도심지터널" AND 인접구조물="유" AND 토피고="2D이하" AND RMR="40-21, 20-0"	강관다단그라우팅 또는 대구경 강관다단그라우팅
28	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="100-41" AND 단층="무"	Fore-poling
29	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="100-41" AND 단층="유(통과)"	Fore-poling
30	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="100-41" AND 단층="유(2D이하)"	Fore-poling
39	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="40-21" AND 단층="무" AND 인접구조물="무"	Fore-poling 또는 강관다단그라우팅
40	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="40-21" AND 단층="무" AND 인접구조물="유"	Fore-poling
32	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="40-21" AND 단층="유(통과, 2D이하)"	Fore-poling
41	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="20-0" AND 단층="무" AND 인접구조물="무"	Fore-poling
42	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="20-0" AND 단층="무" AND 인접구조물="유"	대구경 강관다단그라우팅
34	시공조건="산악터널" AND 갱구부="본선구간" AND RMR="20-0" AND 단층="유(통과)"	Fore-poling
35	시공조건="산악터널" AND 갱구부="갱구부" AND RMR="100-41" AND 토피고="2D이상"	Fore-poling
36	시공조건="산악터널" AND 갱구부="갱구부" AND RMR="100-41" AND 토피고="2D이하"	강관다단그라우팅
16	시공조건="산악터널" AND 갱구부="갱구부" AND RMR="40-21"	강관다단그라우팅
17	시공조건="산악터널" AND 갱구부="갱구부" AND RMR="20-0"	강관다단그라우팅, 필요시 대구경강관다단그라우팅

의사결정트리 분석에 의한 추론규칙과 분류의 타당성을 검증하기 위해서 오분류확률을 <표 5>와 같이 나타내었다. 의사결정트리 분석시 정의된 유의확률( $p < 0.15$ )이 유의할지라도 향후 예측을 위한 모델을 구축하기 위해서 오분류확률과 표준오차를 확인하였다. 오분류확률을 보면 33.11%로서  $(6+6+1)+(9+10+1)+(3+6+0)+(1+0+7)=50$ 에 대하여  $50/151 \times 100 = 33.11\%$ 의 계산과정을 거친다. 그리고 목표변수가 얼마나 정확하게 분류되었는가를 실제범주와 예측범주의 패턴을 통해 알 수 있는데, <표 5>에서 음영으로 표시된 부분은 해당 보조공법이 예측과 분류시 실제 데이터와 일치한 것이고 음영 이외의 부분은 일치하지 않는 것을 나타낸다. 따라서 보조공법에 대한 의사결정트리의 전체적인 분류의 정확도는  $33.11\% \pm 3.82\%$ 으로 약 70% 내외로 판명된다.

표 5. 보조공법의 오분류확률

실제 \ 예측	Fore-poling	강관다단 보강그라우팅	대구경 강관다단 보강그라우팅	조합공법	계
Fore-poling	48	9	3	1	61
강관다단 보강그라우팅	6	32	6	0	44
대구경 강관다단 보강그라우팅	6	10	19	7	42
조합공법	1	1	0	2	4
계	61	52	28	10	151
오분류확률		0.331126			
표준오차		0.038298			

#### 4. PDA를 이용한 보조공법의 의사결정

PDA는 정보의 수집, 저장, 작성, 검색 및 통신기능이 결합된 휴대를 목적으로 하는 정보 단말기로서 이동성과 편의성을 제공하는데, 이를 터널현장에서 직접 사용할 수 있도록 의사결정트리 분석을 통해 추론된 규칙을 기반으로 PDA S/W를 구축하였다. PDA S/W를 이용하여 적절한 터널의 보조공법을 제공함으로써 터널기술자는 시공 중 발생하는 불확실한 상황에 능동적으로 대처하고 의사결정 과정에 걸리는 시간을 단축할 수 있을 것이다. 본 연구에서 구축한 PDA S/W는 Embedded Visual Basic 3.0으로 개발하였으며, 시스템 개발시 PDA와 PC에 설치한 프로그램 및 장비는 <표 6>과 같다.

표 6. 개발장비 및 프로그램

구분	종류 및 제원
개발언어	Microsoft Embedded Basic 3.0
SDK	Microsoft Pocket PC 2002 SDK
Platform	Microsoft Window CE Platform Manager 4.0
Sync.	Microsoft Active Sync. 3.7 이상 버전
PDA	PDA (hp IPAQ Pocket PC h2210)

PDA S/W는 사용자가 입력변수를 선택하여 입력함으로서 해당하는 보조공법이 제시될 수 있도록 구성되었으며, 도심지터널과 산악터널로 구분하였다. <그림 4>는 도심지터널에서 인접구조물, RMR, 단층, 토피고에 해당하는 항목을 선택하여 “강관다단그라우팅 또는 대구경강관다단 그라우팅”의 보조공법을 제시하였으며, 산악터널에서는 갱구부, RMR, 단층, 토피고, 인접구조물을 각각 선택하여 “Fore-poling”의 보조공법을 제시하였다.





그림 4. 도심지터널과 산악터널에서 보조공법 결정

## 5. 결론

터널의 보조공법은 지반의 불확실성과 급변할 수 있는 시공환경 등 불리한 여건에서 적용되기 때문에 현장상황에 따라 효율적인 의사결정이 수행되어야 한다. 본 연구에서는 국내 터널현장에서 적용된 보조공법을 바탕으로 데이터마이닝 기법인 의사결정트리를 적용하여 설계변수가 보조공법에 미치는 교호효과를 분석하였다. 분석결과를 통한 결론은 다음과 같다.

1. 터널의 보조공법에 가장 큰 영향을 미치는 설계변수는 시공조건으로 도심지터널과 산악터널에 적용되는 보조공법이 상이한 것으로 나타났다.
2. 도심지터널은 낮은 토피고와 주변 밀집 건물, 도로, 교량 등의 영향으로 주로 강관다단보강 그라우팅공법과 대구경 강관다단보강 그라우팅공법이 적용되었다.
3. 산악터널은 갱구부의 영향에 따라 보조공법이 적용되었으며, 갱구부 구간은 RMR에 따라 보조공법이 결정되었다. 반면 본선구간은 RMR, 단층, 인접구조물에 따라 보조공법이 결정되었다.
4. 의사결정트리 분석을 통해 설계변수간 규칙을 추론하여 PDA S/W를 구축하였다.

## 참고문헌

1. Biggs, D., B. de Ville and E. Ville, A method of choosing multiway partitions for classification and decision tree, Journal of Applied Statistics, Vol. 18, pp.46-62, 1991
2. Breiman, L., J. H. Friedman, R. A. Olshen and C. J. Stone., Classification and regression trees, Wadsworth, Belmont, 1984
3. Loh, W., Y. Shih. Forthcoming, Split selection Methods for Classification trees, Statistics Sinica, Taiwan, 1997
4. Kass, G., An exploratory technique for investigating large quantities of categorical data, Applies Statistics, 1980
5. Quinlan, J. R., C4.5 Programs for machine learning, Sna Mateo : Morgan Kaufmann, 1993
6. 강현철, 한상태, 최종후, 김은석, 김미경, “데이터마이닝 방법론 및 활용”, 자유아카데미, 2001
7. 한국철도시설공단, “철도설계편람, 터널편 (VII)”, 2004