

모발분석 결과를 이용한 대사형의 분류 *

한근식¹⁾

요 약

모발분석(TMA: Tissue Mineral Analysis)은 머리카락 속에 있는 30여 가지의 미네랄과 8가지의 중금속의 양과 중요 미네랄 비율을 분석하여 체내에 과잉, 결핍 및 불균형 상태를 평가할 뿐만 아니라 체내에서 미네랄들의 상호연관 관계를 평가하여 건강을 유지하는 방향을 제시하는 모발조직 검사방법을 말한다. 현재 세계 46개국 의료기관에서 임상을 위해 널리 사용중이나 국내에는 아직 인지도가 낮아 널리 활용되지 못하고 있으나 점차로 국내 대학병원을 중심으로 필요성을 인식하고 있다. 본 연구에서는 한국티아이로 부터 제공받은 1,000명의 사례자료를 이용하여 8가지 대사형(Metabolic Type)을 결정트리분류기로 분류하였다.

주요용어: 모발분석, 결정트리, 대사형.

1. 대사형의 분류

모발조직검사는 조직 중 미네랄의 결핍, 과잉 및 불균형을 평가하는 방법이다. 이 검사는 영양, 미량 및 독성 미네랄의 정상, 과잉, 결핍 및 불균형을 제시하고 자율신경 Dominance에 의한 대사형(Metabolic type)을 제시하며 호르몬의 세포수준활성을 설명하고 개인에 알맞은 식이추천을 제공한다. 특히 모발조직의 미네랄 농도는 혈액이나 다른 조직에서 발견되는 미네랄보다 농도가 매우 높아 중금속 오염의 확인, 마약성분 검출에 널리 이용되고 있다. 본 연구에서는 모발분석에 참가한 1,000여명의 한국인을 대상으로 미네랄과 대사형과의 관계를 분석하는데 초점을 맞추고 있다.

대사형은 우리가 일반적으로 알고 있는 태양인, 태음인, 소양인, 소음인 등과 같은 체질의 서구식 분류방법으로 이 분류에 의한 체질개선으로 질환을 예방하고 체내 오염을 치유하는 예방의학이 널리 보급되고 있다.

본 연구에서 지능적 추론의 배경이 되는 대사형의 분류는 결정트리분류기(Decision Tree Classifier)를 이용하며 이를 통하여 인체의 8가지 대사형(Fast1 : F_1 , Fast2 : F_2 , Fast3 : F_3 , Fast4 : F_4 , Slow1 : S_1 , Slow2 : S_2 , Slow3 : S_3 , Slow4 : S_4)을 분류하고자 한다. 1단계 분류기는 클래스간의 변별력 중에서 평균이 최대가 되는 특징조합을 이용하나, 이 특징조합이 모든 클래스에 대하여 최적의 변별력을 항상 갖는 것이 아니므로 두 종류이상의 결정을 적절한 수순에 따라 몇 번이고 반복하여 입력패턴의 클래스를 결정하는 다단계분류기(Safavian과 Landgrebe, 1991, Goodman과 Smith, 1988)인 결정트리 분류방법을 응용하여 분류한다.

* 이 연구는 2003년도 한신대학교 특별연구비에 의하여 연구되었음.

1) (447-791) 경기도 오산시 양산동 411, 한신대학교 정보과학대학 정보시스템학과, 교수

E-mail: gshan@hanshin.ac.kr

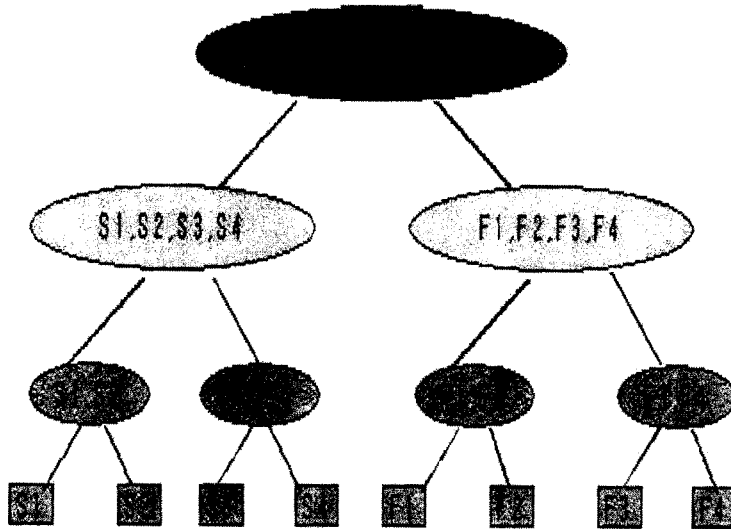


그림 1.1: 결정트리에 의한 대사형의 분류

2. 결정트리분류기에 의한 분류방법

위의 그림 1.1에서 결정트리분류기가 패턴집합을 분류하는 과정은 다음과 같다. 우선 루트노드에서 영양미네랄인 Ca 와 P 의 비, Ca/P 를 이용하여 느린 대사타입(Slow Metabolic Type)의 패턴집합 $\{S_1, S_2, S_3, S_4\}$ 와 빠른 대사타입(Fast Metabolic Type) 패턴집합 $\{F_1, F_2, F_3, F_4\}$ 의 서브패턴집합으로 분류한 다음, 다시 내부노드에서 Ca/K 를 이용하여 1차 분류된 서브패턴 집합 $\{S_1, S_2, S_3, S_4\}$ 를 패턴 $\{S_1, S_2\}$ 와 $\{S_3, S_4\}$ 로 분류하고 Na/Mg 를 이용하여 1차 분류된 서브패턴 집합 $\{F_1, F_2, F_3, F_4\}$ 를 $\{F_1, F_2\}$ 와 $\{F_3, F_4\}$ 로 분류하였다. 마지막 단계에서 Na/Mg 를 이용하여 2차 분류된 패턴집합 $\{S_1, S_2\}$ 와 $\{S_3, S_4\}$ 를 S_1 과 S_2 그리고 S_3 과 S_4 로 최종 분류하고, 같은 방법으로 Ca/K 를 이용하여 2차 분류된 패턴집합 $\{F_1, F_2\}$ 와 $\{F_3, F_4\}$ 를 F_1 과 F_2 그리고 F_3 와 F_4 로 최종 분류하여 결정트리를 구성하였다.

대사형의 분류는 특성변수에 대한 클래스별 분포를 이용해도 같은 결과를 보여 준다. 다음은 특성변수를 계급적(hierarchical)으로 적용하여 클래스별 분포를 이용한 대사형의 분류를 보여준다.

우선 Ca/P 클래스 분포에 의해 대사타입을 $\{S_1, S_2, S_3, S_4\}$ 와 $\{F_1, F_2, F_3, F_4\}$ 의 서브패턴집합으로 분류한 것이 그림 2.1이다. 이 그림에서 왼쪽의 4개 분포는 침도가 높고 표준편차가 작은 대사형 F 의 Ca/P 분포로 오른쪽부터 $F_2 \sim (\hat{\mu} = 1.85, \hat{\sigma} = 0.28)$, $F_1 \sim (\hat{\mu} = 1.97, \hat{\sigma} = 0.38)$, $F_3 \sim (\hat{\mu} = 2.01, \hat{\sigma} = 0.32)$, $F_4 \sim (\hat{\mu} = 2.34, \hat{\sigma} = 0.26)$ 를 보여주고 있으며 아래쪽으로 누운 형태의 4개 분포는 침도가 낮고 표준편차가 큰 대사형 S 의 Ca/P 분포로 오른쪽부터 $S_4 \sim (\hat{\mu} = 3.79, \hat{\sigma} = 1.64)$, $S_2 \sim (\hat{\mu} = 5.35, \hat{\sigma} = 3.45)$, $S_3 \sim (\hat{\mu} = 5.75, \hat{\sigma} = 3.60)$, $S_1 \sim (\hat{\mu} = 9.31, \hat{\sigma} = 6.65)$ 을 보여주고 있어 Ca/P 의 분포에 의해 대사형 F 와 S 가 분류됨.

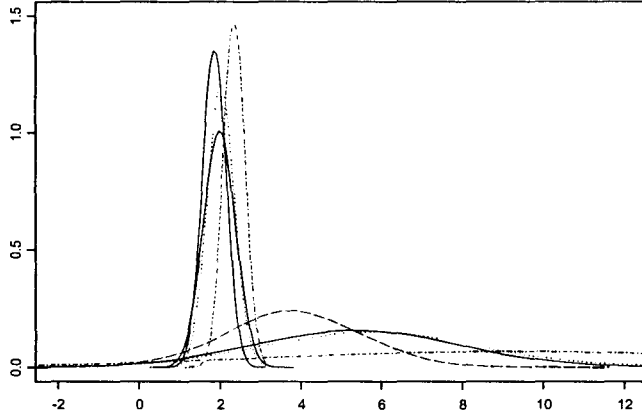


그림 2.1: Ca/P 클래스 분포를 이용한 대사타입의 분류

볼 수 있다.

위와 같이 분류된 대사형 S 를 Ca/K 클래스 분포에 의해 분류하면 다음 그림 2.2와 같다. 그림을 살펴보면 Ca/K 의 분포는 첨도가 높고 표준편차가 작은 $S_4 \sim (\hat{\mu} = 1.78, \hat{\sigma} = 1.15)$,

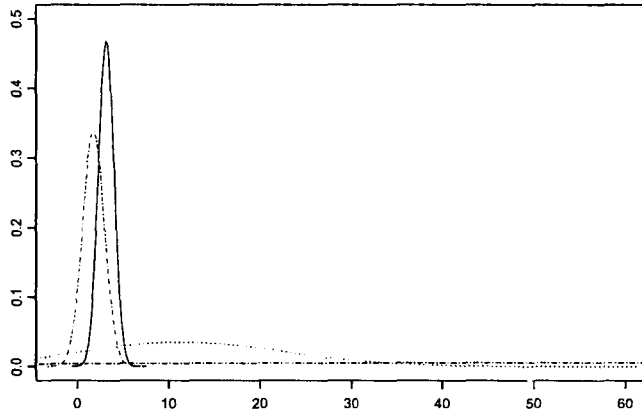


그림 2.2: 대사형 S 의 Ca/K 클래스 분포에 의한 분류

$S_3 \sim (\hat{\mu} = 3.16, \hat{\sigma} = 0.81)$ 가 왼쪽에 위치하고 아래쪽에 첨도가 낮고 표준편차가 큰 $S_2 \sim (\hat{\mu} = 11.7, \hat{\sigma} = 11.03)$, $S_1 \sim (\hat{\mu} = 54.79, \hat{\sigma} = 71.8)$ 로 분류된다.

이제 Ca/K 클래스 분포에 의해 분류된 S_3 와 S_4 그리고 S_1 과 S_2 를 Na/Mg 클래스 분포에 의해 분류하면 그림 2.3, 그림 2.4와 같다. 그림 2.3에서 분포를 보면 왼쪽의 높은 첨도와 낮은

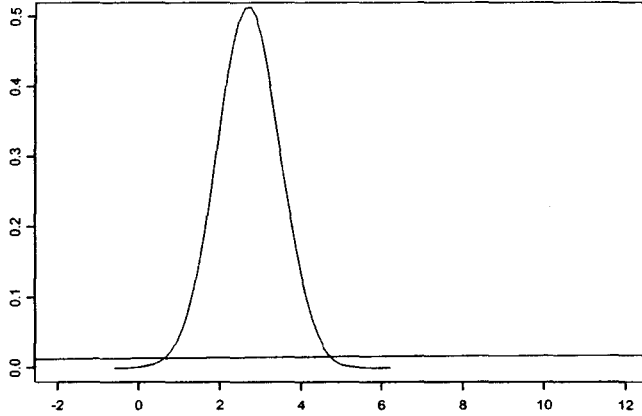


그림 2.3: Na/Mg 클래스 분포에 의한 S_3 와 S_4 의 분류

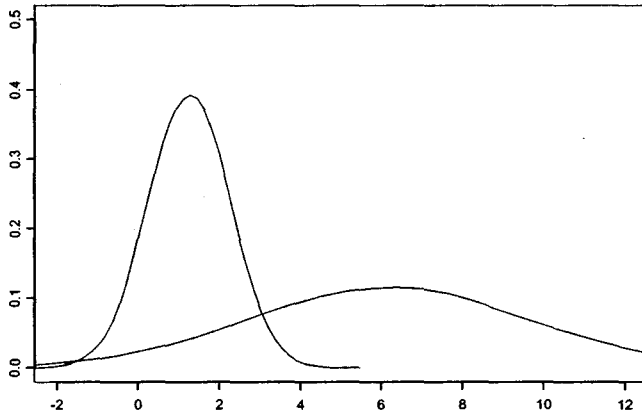


그림 2.4: Na/Mg 클래스 분포에 의한 S_1 과 S_2 의 분류

표준편차를 갖는 분포는 $S_3 \sim (\hat{\mu} = 2.7, \hat{\sigma} = 0.76)$ 이고 아래쪽의 분포는 표본편차가 매우 큰 $S_4 \sim (\hat{\mu} = 15.3, \hat{\sigma} = 21.27)$ 로 분류됨을 알 수 있다. 마찬가지로 그림 2.4에서 Na/Mg 클래스 분포를 보면 왼쪽의 높은 첨도와 낮은 표준편차를 갖는 분포는 $S_2 \sim (\hat{\mu} = 1.25, \hat{\sigma} = 0.97)$ 이고 오른쪽의 분포는 표본편차가 크고 첨도도 낮은 $S_1 \sim (\hat{\mu} = 6.16, \hat{\sigma} = 3.33)$ 으로 분류됨을 알 수 있다.

한편, 대사타입 F 의 Na/Mg 클래스 분포에 의한 분류는 그림 2.5와 같다. 위 그림에서 왼쪽의 첨도가 높은 분포가 $F_4 \sim (\hat{\mu} = 1.78, \hat{\sigma} = 0.95)$ 와 $F_3 \sim (\hat{\mu} = 2.62, \hat{\sigma} = 0.75)$ 이고 표준편차가 큰 아래쪽의 분포는 $F_2 \sim (\hat{\mu} = 6.04, \hat{\sigma} = 3.07)$ 와 $F_1 \sim (\hat{\mu} = 29.6, \hat{\sigma} = 27.33)$ 이다.

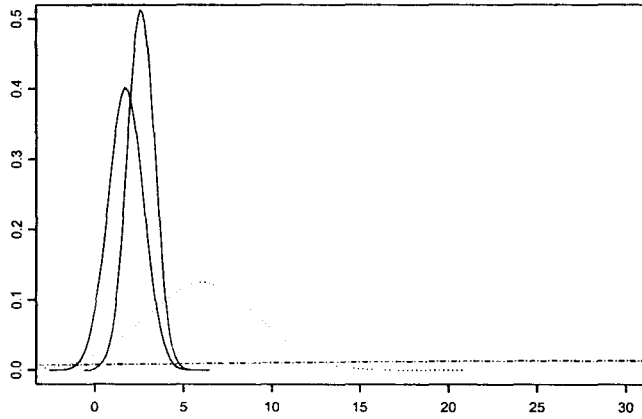


그림 2.5: 대사타입 F 의 Na/Mg 클래스 분포에 의한 분류

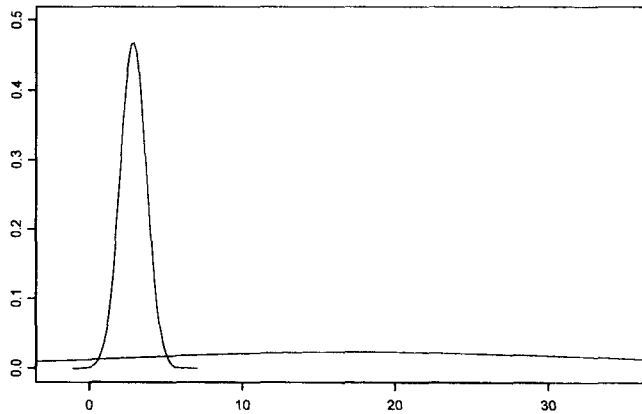


그림 2.6: Ca/K 클래스 분포에 의한 F_3 와 F_4 의 분류

끝으로 침도가 높은 F_3 와 F_4 는 Ca/K 클래스에 의한 분포로 그림 2.6과 같이 분류되고 F_1 과 F_2 는 Ca/K 클래스에 의해 그림 2.7과 같이 분류된다. TMA에서 가장 기본이 되는 대사형의 분류 후에 각 대사형 별 영양미네랄과 독성미네랄, 영양미네랄들의 비(ratio), 독성미네랄들의 비등과 영양상태와의 관계는 상관분석(correlation analysis)을 통해 상관계수가 0.3보다 큰 미네랄들을 대상으로 분석하였다.

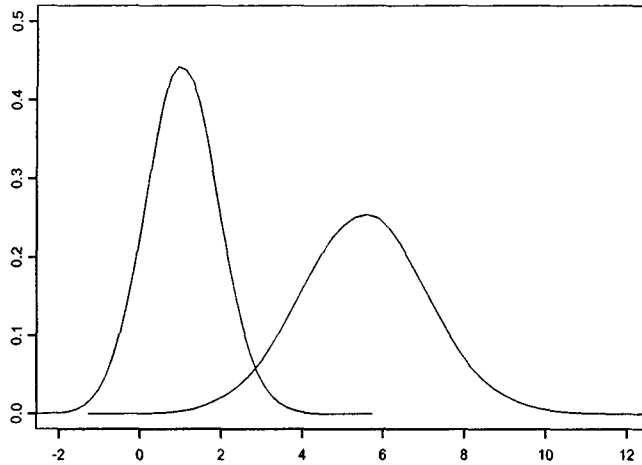


그림 2.7: Ca/K클래스 분포에 의한 F_1 과 F_2 의 분류

3. 통계적 분류의 적절성

분류방법의 적절성을 확인하는 절차중의 하나는 오분류율(misclassification probability)을 계산하는 것이다. 오분류율의 계산은 표본을 이용하여 다음과 같은 식 (3.1)에 의해 계산한다.

$$ARE(actual\ error\ rate) = p_1 \int_{\hat{R}_2} f_1(x)dx + p_2 \int_{\hat{R}_1} f_2(x)dx \quad (3.1)$$

여기에서 \hat{R}_1 과 \hat{R}_2 는 표본의 크기에 의해 결정되는 분류지역(classification regions)을 의미한다.

그러나 위의 방법을 이용하기 위해서는 식 (3.1)에서 보는 바와 같이 미지의 밀도함수(density function), $f_1(x)$ 와 $f_2(x)$ 를 알아야 한다. 표본정보만을 갖고 있는 본 연구에서는 모집단의 분포에 종속되지 않고 오분류율을 계산할 수 있어야 하기 때문에 Apparent error

표 3.1: 오분류율 측정방법

		Predicted membership		
		π_1	π_2	
Actual membership	π_1	n_{1C}	n_{1M} $= n_1 - n_{1C}$	n_1
	π_2	n_{2M} $= n_2 - n_{2C}$	n_2	n_2

rate(APER: Johnson과 Wichern, 1992)을 이용하였다.

위 표 3.1을 이용하여 $APER = \frac{n_{1M} + n_{2M}}{n_1 + n_2}$ 로 계산되며 이는 오분류된 비율을 나타낸다. 여기에서 n_{1C} = 실제 π_1 에 속하면서 π_1 에 분류된 수, n_{1M} = 실제 π_1 에 속하면서 π_2 로 잘못 분류된 수, n_{2C} = 실제 π_2 에 속하면서 π_2 에 분류된 수, n_{2M} = 실제 π_2 에 속하면서 π_1 으로 잘못 분류된 수이다.

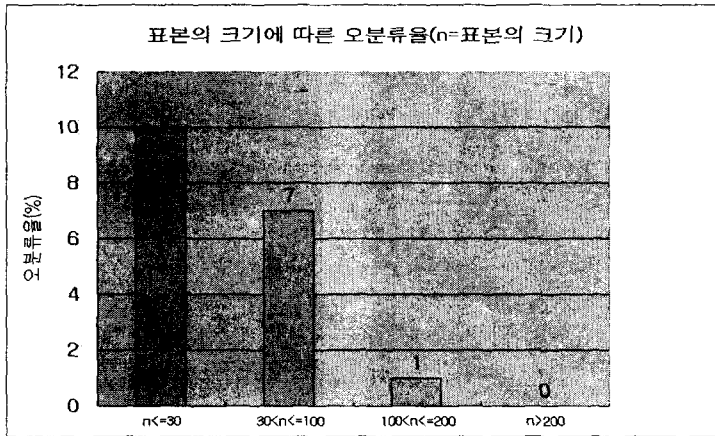


그림 3.1: 오분류율

본 논문에서 오분류율 측정 결과 8가지 대사형(빠른대사, 느린대사)에 대해서는 표본의 크기가 충분히 커서 오분류율이 거의 0에 가까웠으나 표본의 크기가 작은 미네랄간의 관계 등에서는 다음 분포와 같은 오분류가 발생하였음을 알 수 있었다.

참고문헌

Safavian, S.R. and Landgrebe, D. (1991). A survey of decision tree classifier methodology, *IEEE Trans. on Syst., Man, and Cybern.*, **21**, 660-674.

Brown, D.E., Corruble, V. and Pittard, C.L. (1993). A Comparison of decision tree classifiers with backpropagation neural networks for multinomial classification problems, *Pattern Recognition*, **26**, 953-961.

Strain, W.H., Pories, W.J., Flynn, A. and Hill, O.A. (1972). *Trace Element Nutriture and Metabolism Through Head Hair Analysis*. Trace Substances in Environmental Health. Hemphill, D., Ed. Univ. Mo. Press, Columbia.

Laker, M. (1982). *On Determining Trace Element Levels in Man The Uses of Blood and Hair*. Lancet 11.

Johnson, R.A. and Wichern, D.W. (1992). *Applied Multivariate Statistical Analysis*. Prentice Hall.

The Classification of Metabolic Type Using Tissue Mineral Analysis *

Geunshik Han ¹⁾

ABSTRACT

We have sent the 1000 hair samples to USA and received the results from USA because the programs for the interpretation of Tissue Mineral Analysis (TMA) result is not opened yet in Korea. Therefore, the study will analyze the relationship between the hair for Korean and minerals and make a classification system. To achieve the goal, first of all, we coded the results of patients and classified the analyzed results which are the interrelationship between the minerals and dietary situation and the heavy metals and diseases through the statistical methods. Finally we classified 8 metabolic type using decision tree classifier.

Keywords: Hair tissue analysis, Decision tree, Metabolic type

* This work was supported by grant from Hanshin university.

1) Professor, Department of Information System Engineering, Hanshin university, 411 Yangsan-dong, Osan, Kyunggi-do, 447-791, Korea.

E-mail: gshan@hanshin.ac.kr