14강. 분류분석(2)

◈ 담당교수 : 이윤동 교수

■ 주요용어

용어	해설
불순도척도	나무방법에서 가지를 분할해가는 기준으로, 일종의
	적합결여도에 해당됨.
복잡도기준	나무방법에서, 말단노드의 수를 제약하기 위하여 불
	순도척도와 말단노스의 수를 동시에 고려하여 만든
	척도
서포트벡터	전체자료 중, 경계식을 만드는데 직접 관여하는 자
	료.
편차변수	서포트벡터기계에 의한 분류에서 오분류된 양을 의미
	하는 값으로, 일종의 적합결여 정도를 의미함.

■ 연습문제

1. 나무방법을 이용할 수 있는 대표적인 분야는 무엇인가?

정답 및 해설 : 회귀분석과 판별분석

2. 분류분석목적의 나무방법에서 사용할 수 있는 불순도척도로는 어떤 것들이 있는 가?

정답 및 해설: 오분류 오차, 지니계수, 크로스 엔트로피

3. 서포트벡터기계 방법에서 자주이용되는 재생커널은 어떤 종류들이 있는가?

정답 및 해설: linear, polynomial, radial, signal.

■ 참고자료

- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of ideas immanent in neural activity.
 - Bulletin of Mathematical Biophysics 5, 115-133.
- Rao, R. C. (1948). The utilization of multiple measurements in problems of biological classification.
 Journal of the Royal Statistical Society, Series B 10 (2): 159-203.
- Ripley, B. D. (1993) Statistical aspects of neural networks.
 In Networks and Chaos Statistical and Probabilistic Aspects, eds
 O. E. Barndorff-Nielsen, J. L. Jensen and W. S. Kendall, pp. 40-123.
 London: Chapman & Hall.
- Ripley, B. D. (1994) Neural networks and flexible regression and discrimination.
 - In Statistics and Images 2, ed. K. V. Mardia, volume 2 of Advances in Applied Statistics, pp. 39-57. Abingdon: Carfax.
- Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with S. Springer Science & Business Media.

■ 코스웨어

- R의 기본 코스웨어 학습방법 안내

다음은 R에 대한 기본적인 학습을 할 수 있는 방법을 안내 해 드립니다. 아직 R에 대하여 익숙하지 못한 학습자는 "R의 입문 코스웨어"를 통하여 R에 대하여 익숙해지기 바랍니다.

1) [학교 홈페이지에 접속] - [우측하단"학과"선택]



2) ["정보통계학과" 선택]



3) [상단메뉴 학습정보 내의 "학습자료실" 선택]



4) [좌측메뉴의 "코스웨어" 선택]



5) [좌측메뉴의 "R의 입문 코스웨어" 선택]



6) "R의 입문 코스웨어"는 총10장으로 구성됨



7) 각 장의 좌측메뉴의 학습목차를 참고하여 학습



"R의 입문 코스웨어" 과정을 통하여 R의 기본개념 및 구조를 이해하고 실제 작업에 필요한 자료처리 및 분석기법을 응용할 수 있는 능력을 학습하기 바랍니다.