

베이지안 네트워크와 방사형 그래프를 이용한 섬망의 효과 규명

이제영¹ · 배재영²

¹²영남대학교 통계학과

접수 2011년 8월 8일, 수정 2011년 8월 25일, 게재확정 2011년 9월 5일

요약

최근 의학에서는 정신 질환과 관련된 위험 인자를 찾는 것이 중요해지고 있다. 인자들을 찾아서 인자들의 특성과 관련성을 파악하면 병을 사전에 예방 할 수 있다. 또한 이 연구는 의학 발전에 많은 도움을 줄 수 있다. 정신 질환에 대한 위험요인은 주로 로지스틱 회귀모형을 사용하여 찾아 왔다. 하지만 이 논문에서는 데이터마이닝 기법 중 CART, C5.0, 로지스틱, 신경망, 베이지안 네트워크 방법을 이용한다. 정신장애 질병인 섬망자료를 적용하여, 최적의 모형인 베이지안 네트워크 방법을 선택 하였다. 이 베이지안 네트워크 기법을 위험 요소를 찾는 데 사용하고, 이 위험인자 간의 관계를 방사형 그래프를 통해서 규명하였다.

주요용어: 데이터마이닝, 베이지안 네트워크, 정신장애 질병 섬망.

1. 서론

최근 우리나라에서는 치매 (dementia)와 우울증과 섬망 (delirium)과 같은 정신장애와 관련이 있는 질병들이 증가하고 있는 추세이다. 그래서 많은 사람들이 정신장애 및 질병에 대한 관심도 점차 높아지고 있다. 여기서 우리는 정신장애 질병 중에서도 최근 치매와 함께 주목받고 있는 섬망이라는 질병에 관심을 가졌다. 섬망은 혼돈과 비슷하지만 심한 과다행동과 생생한 환각, 초조함과 떨림이 있는 것으로, 의식장애와 내적인 흥분의 표현으로 볼 수 있는 운동성 흥분을 나타내는 병적 정신상태를 말한다.

특히 고령에 뇌수술을 받은 환자의 경우 섬망을 호소하는 사례가 빈번하게 나타날 수 있다고 한다. 섬망은 여러 가지 수술로 인해서 입원 치료중인 70세 이상 노인환자의 30%에서 나타날 정도로 흔한 정신장애 질환이다 (Inouye, 1994; Cole과 Primeau, 1993; Inouye 등, 1999). 현재까지 섬망에 대한 연구는 주로 외국에서 이루어졌다. 주요 연구들을 살펴보면, 중환자실에 입원해 있는 환자들의 섬망관련 위험인자에 대한 연구 (Dubois 등, 2001), 중환자실에 입원해 있는 환자들의 섬망관련 32가지 위험인자 연구 (Arend와 Christensen, 2009) 등이 있다. 그러나 중환자실에 입원해 있는 환자들만 연구를 하여서 응급실을 통한 환자의 섬망 위험인자 연구는 많이 부족한 실정이다.

섬망과 비슷한 정신관련 질병의 위험인자들을 찾는 방법으로 로지스틱 회귀모형이 주로 이용되었다 (Kwak 등, 2011). 또한 정신관련 질병 이외의 병에서도 로지스틱 회귀모형이 많이 이용되었다 (강석복 등, 2010). 로지스틱 회귀모형이 정신장애 질병과 관련된 위험인자를 쉽고 간단하게 찾을 수 있지만 복잡한 상황에 유연하게 대처하지 못한다는 단점이 있다. 이는 위험인자 간에 상호작용을 얻고 싶을

¹ (712-749) 교신저자: 경북 경산시 대동 214-1 영남대학교 통계학과, 교수. E-mail: jlee@yu.ac.kr

² (712-749) 경북 경산시 대동 214-1 영남대학교 통계학과, 대학원, 석사과정.

때, 의미 있는 위험인자들을 충분히 얻지 못하는 어려움이 있기 때문이다. 이 논문에서는 이러한 점을 보완하고, 좀 더 다양한 데이터마이닝 기법을 사용하여서 섬망 자료에 적용을 한다. 여기서 데이터마이닝이란 데이터 웨어하우스 (data warehouse)와 같은 자료저장소에 저장되어 있는 방대한 양의 데이터로부터 통계분석과 모델링을 통해서 의사결정에 도움이 되는 유용한 정보를 찾아내는 것이다 (Tan 등, 2007).

본 논문 2절에서는 데이터마이닝 기법인 로지스틱 회귀모형 (logistic regression), C5.0, 신경망 (neural network), CART (classification and regression tree), 베이저안 네트워크 (Bayesian network)를 소개하고, 3절에서는 섬망자료 소개와 이 자료를 데이터마이닝 기법에 적용하여 최적의 모형을 선별한다. 4절에서는 최적의 모형으로 섬망과 관련된 주요 위험인자를 선별한 뒤 이 인자들 간의 관계를 방사형으로 규명한다.

2. 데이터마이닝 기법 소개

2장에서는 데이터마이닝 기법인 CART, C5.0, 로지스틱 회귀모형, 신경망, 베이저안 네트워크기법에 대해서 간략하게 소개를 한다 (이용원, 2004).

CART는 각 (독립)변수를 이분화 (binary split)하는 과정을 반복하여 트리 형태를 형성함으로써 분류 (종속변수가 범주형일 때)와 회귀분석 (종속변수가 연속형일 때)을 수행하는 의사결정나무 방법 중의 하나이다 (Breiman 등, 1984). CART는 C5.0과 다르게 목표변수가 연속형인 경우에도 활용을 할 수 있는 것이 특징중의 하나이다. CART는 지니 지수를 이용하여 데이터의 불순도를 측정한다. 의사결정나무인 CART는 지니 측도에 의해서 나무형 처럼 가지가 나누어지게 되며, 지니 계수를 기준으로 가장 성취도가 좋은 변수 및 수준을 찾는다 (허명회와 이용구, 2008).

C5.0은 Ross Quilan의 C4.5를 개량한 알고리즘으로 의사결정 나무 모델의 하나이며, 범주형 (flag / set type)인 목표변수를 이용하여 다른 입력변수들의 분류를 통해, 세분화 모델이나 목표변수를 예측할 수 있는 분석 방법이다. 이 C5.0에서는 주로 사용 되는 기본 개념인 무질서를 측정하는 하나의 측정 지표가 엔트로피이다. 무질서와 확률을 이용한 개념이 데이터에서 구분을 잘 짓고, 못 짓는 속성 값을 찾아내는 지수로 사용되는 것으로 의사결정나무에서 활용을 한다. 의사결정나무인 C5.0은 이 엔트로피를 기준으로 가장 엔트로피를 낮추는 분리 변수를 찾고자 하는 것이다 (Beom과 Song, 2009).

로지스틱 회귀모형은 단지 두 개의 값만을 가지는 목적변수와 설명변수들 사이의 인과관계를 분석하는 하나의 통계기법이다. 또한 병리학 연구에서 질병과 연관 있는 위험요인들을 식별하거나, 임상 연구 자료에 중요한 요인들을 식별하는 탐색적 분석에 많이 적용된다 (이재원 등, 2005; 허명회와 이용구, 2008). 일반적인 회귀분석과 유사한데, 종속변수가 구간 비율척도 변수가 아니라 0,1의 값을 가지고 이분형 범주형 변수라는 것이 다른 점이다. 또한 회귀계수로 독립변수들의 오즈비를 쉽게 구할 수 있는 장점이 있다.

신경망은 인간 두뇌의 신경망을 흉내 내어 실제 자신이 가진 데이터로부터의 반복적인 학습 과정을 거쳐 데이터에 숨어 있는 패턴을 찾아내는 모델링 기법이다 (Warren, 1994). 이 신경망 분석은 데이터 마이닝 기법 중 가장 많이 알려진 것이며, 보통 데이터마이닝에서 신경망 분석을 통한 패턴을 찾아내는 것을 연상할 만큼 잘 알려진 분석이다. 특징으로는 인간이 학습하는 구조를 응용하여 자료 분석에 이용을 하며, 자료의 패턴이 변화함에 따라 이를 학습하고, 이에 가중치를 변화 적용하여, 최적의 해를 구하는 것이다. 통계적 가정 등에서 자유롭고, 복잡한 상호작용을 가진 자료에 탁월한 능력을 발휘 한다. 반면에 결과를 해석하는데 어려움이 있어서 적절하게 활용하는 것이 쉽지 않은 단점이 있다 (Min과 Hyun, 2009).

베이저안 네트워크는 변수들의 집합 간의 관계를 확률을 이용하여 그래픽 모델로 나타낸 기법이다

(Tan 등, 2007). 연구 분야의 수치적인 지식을 확률적으로 표현할 수 있는 방법 중의 하나로써, 변수들의 조건부 관계를 확률적으로 나타내며, 변수간의 관계를 네트워크를 통해서 쉽게 나타낼 수 있다. 과거에 베이저안 네트워크는 인코딩이 불확실한 전문적인 지식에 가장 유용한 기법으로 사용되어 왔다 (Heckerman, 1995; Jensen, 1996). 이 베이저안 네트워크는 변수 집합 사이의 의존 관계를 표현하기 위해서 방향성 비순환 그래프 (Directed Acyclic Graph : DAG)와 각 노드를 앞의 노드들과 연관시키는 확률표로 정의 할 수 있다. 그림 2.1는 방향성 비순환 그래프를 나타내는 예로써 확률변수 A, B, C의 관계를 네트워크로 나타낸 것이다.

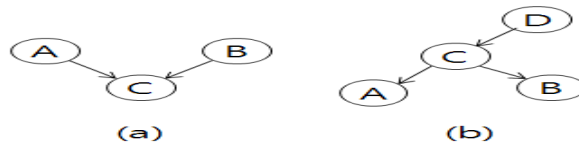


그림 2.1 방향성 비순환 그래프를 사용한 확률적 관련성 표현

그림 2.1를 살펴보면 노드로 되어 있는 것이 변수가 되겠고, 화살표가 변수들의 종속적인 관계를 보여주는 것을 확인 할 수 있다. (a)의 그림을 살펴보면 A, B가 변수 C에 영향을 주고 있는 것을 확인할 수 있다. 이는 A, B는 독립변수이고 C는 목표변수라고 할 수 있다. (b)의 그림을 살펴보면 C가 A, B로 화살표가 있는 것을 확인할 수 있다. 이는 C가 A, B의 부모노드라고 할 수 있으며, C에 화살표를 보내는 D는 C의 부모노드이고, A, B의 조상노드라고 할 수 있다. 베이저안 네트워크는 어떤 노드가 부모노드로 알려져 있다면 그 노드는 자손이 아닌 노드들과 조건부로 독립적이다. 어느 사상에 대한 발생이 주어진 경우에 특정한 가설의 사실이 조건부 확률을 구하는 방법을 말하는데, 즉 사전에 알고 있는 정보를 염두에 두고 특정 사건이 일어날 확률이다. 베이저안 네트워크는 조건부 확률 계산에 베이즈 정리를 이용하여 다음의 수식을 얻을 수 있다.

$$P(C|A) = \frac{P(A|C)P(C)}{P(A)} \quad (2.1)$$

곱셈규칙을 이용해서 풀면

$$P(C, A) = P(C|A)P(A) = P(A|C)P(C) \quad (2.2)$$

이와 같은 식이 나오며, 체인 룰 (chain rule)을 사용하면

$$P(X_1, X_2, \dots, X_n) = P(X_1) \prod_{i=2}^n P(X_i | X_1, \dots, X_{i-1}) \quad (2.3)$$

식이 만들어진다.

3. 섬망자료와 최종모형 선택

3.1절에서는 이 논문에서 사용될 섬망 자료에 대해서 살펴보고, 3.2절에서는 2절에서 소개된 데이터 마이닝 기법을 활용해 섬망 자료에 적용을 하고, 적용 결과 정확도가 가장 높은 모형을 최종모형으로 선별한다.

3.1. 실험자료

2008년 1월부터 2009년 12월까지 24개월 동안 종합병원 응급실을 통해 내과 병동으로 입원한 70세 이상 고령 환자로 구성되어졌다. 연구에 포함된 환자의 성별, 과거력, 사회력, 생체 징후는 의무기록을 바탕으로 조사된 내용이다. 심망 진단은 정신장애 진단통계 편람 제4편에 의해 정신과적 기저 질환을 제외한 기저 질환에 의해 심망으로 진단받음을 기준으로 하였다 (Kwak 등, 2011). 연구에서 사용한 대상자 수는 414명이며 그 중 42명은 정신과 치료를 통해 기저 질환으로 인한 심망으로 진단받았으며 372명은 진단받지 않았다. 본 논문에서 사용한 위험인자는 2009년 Emma Arend 등이 보고한 32가지 위험인자 (Arend와 Christensen, 2009) 가운데 연령을 제외한 31가지 위험인자를 가지고 심망과 유의한 관련이 있는지를 알아본다. 표 3.1은 분석에 사용된 성별을 제외한 31가지 위험인자를 나타낸 것이다.

표 3.1 심망과 관련된 31가지 위험인자 (Arend와 Christensen, 2009)

Predisposing (질병에 취약한 인자)	Acute illness (급성질병)	Pharmacology (약리학)
Dementia (치매)	Infection (감염)	Anaesthetics (마취약)
Nursing home (요양병원)	Hypoxia (저산소증)	analgesics (진통제)
Alcohol abuse (음주)	Metabolic abnormality (대사이상)	Antibiotics (항생제)
Smoking (흡연)	Electrolyte imbalance (전해질불균형)	anticholinergic (안티콜린제, 농약치료제)
Visualimpairment (시력장애)	Malnutrition (영양실조)	antihistamine (항히스타민제)
Hearing loss (청각장애)	Hemodynamic instability (혈역동학적이상)	Antihypertensives (고혈압약물)
Bun/Cre (신장수치)	Head trauma (두부외상)	bronchodilator (기관지확장제)
Stroke/epilepsy (뇌졸중간질)	Seizure (발작)	cardiacdrug (심장약)
Congestive heart failure (심부전)	Vascular problem (혈관문제)	diuretic (이뇨제)
Depression (우울증)	CNS disorder (central nervous system ; 중추신경계통 이상)	Sedative (진정제)
		Steroid (스테로이드)

3.2. 최종모형 선택

본 논문에서 사용된 심망자료 414명 중에서 42명만이 심망을 앓고 있는 환자들이다. 따라서 심망 환자와 심망이 아닌 환자들의 비율을 맞추기 위해 부스팅을 통하여 심망의 비율을 50:50으로 맞추는 자료를 사용하였다. 또한 훈련자료와 테스트자료의 비율을 각각 70%,30%로 분할하여 분석하였다. IBM SPSS Model 14.1을 이용해서 CART, C5.0, 로지스틱 회귀모형, 신경망, 베이지안 네트워크 기법에 적용하여서, 심망관련 31가지 위험인자를 분석한 결과 나온 정확도를 표 3.2로 나타내었다.

표 3.2 심망자료에 적용된 데이터마이닝 기법들의 정확도 결과

모형	훈련자료 정확도	테스트자료 정확도
CART	0.864	0.775
C5.0	0.81	0.775
Logistic	0.857	0.758
Neural Network	0.867	0.783
Bayesian Network	0.891	0.792

표 3.2에서 나온 훈련자료의 정확도는 CART 모형은 86.4%, C5.0 모형은 81%, 로지스틱 모형은 85.7%, 신경망 모형은 86.7%, 베이지안 네트워크 모형은 89.1%로 나타났다. 훈련자료에서 만들어진 모형을 테스트자료에 적용한 결과는 CART 모형은 77.5%, C5.0 모형은 77.5%, 로지스틱 모형은 75.8%, 신경망 모형은 78.3%, 베이지안 네트워크 모형은 79.2%로 나타났다. 베이지안 네트워크 모형이 다른 모형보다 정확도가 높음을 확인 수 있었다.

훈련자료와 테스트 자료를 종합적으로 봤을 때 정확도가 가장 높게 나타난 베이지안 네트워크를 섬망 자료를 분석할 최종모형으로 선택하였다. 이에 베이지안 네트워크를 섬망 자료에 적용하여 섬망과 관련된 주요 위험인자를 규명하고 섬망과 주요 위험인자 간 관계를 방사형 그래프를 통해 규명한다.

4. 베이지안 네트워크 방법에 의한 주요인자간 방사형 규명

베이지안 네트워크 방법을 섬망자료에 적용한 결과 기존 31가지의 인자중 뇌졸중/간질, 혈액동화적 이상, 이뇨제, 전해질 불균형, 대사이상, 음주, 스테로이드 등 7인자가 섬망에 영향을 많이 미치는 위험인자로 선택되었다. 섬망에 영향을 많이 미치는 위험인자 7가지를 사후확률 값 (섬망 발병률)을 구한 뒤 순서대로 표 4.1에 나타내었다.

표 4.1 7가지 단일위험인자와 사후확률

위험인자	사후확률
뇌졸중/간질	0.881
전해질불균형	0.352
대사이상	0.231
혈액동화적 이상	0.183
이뇨제	0.071
음주	0.044
스테로이드	0.0004

표 4.1은 위험인자가 단일로 존재할 때 사후확률을 얻은 결과이다. 먼저 뇌졸중/간질이 있을 때 섬망이 나타날 확률은 약 88.1%로써 다른 기저질환보다 매우 높은 발병률로 나타났다. 다음으로 전해질 불균형이 있을 때 35.2%의 확률로 섬망이 발병하며, 대사 이상이 있을 때 23.1%로 섬망이 발병한다는 것을 알 수 있다. 혈액동화적 이상이 있는 경우 18.3% 이뇨제를 투약한 경우 7.1%, 음주를 하는 경우가 4.4%로 섬망이 걸릴 확률로 나타났다. 그림 4.1은 사후확률 결과를 섬망에 영향을 많이 미치는 위험인자 순으로 방사형을 구현한 것이다. 어떤 인자가 섬망에 영향을 많이 주는지 한눈에 알아 볼 수 있다.

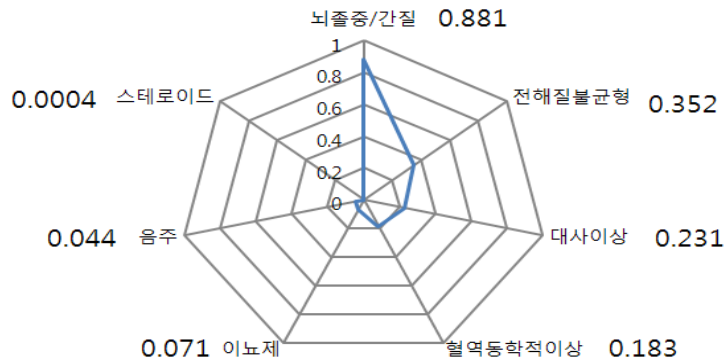


그림 4.1 섬망에 영향을 주는 단일 위험인자를 통한 방사형

그림 4.1은 섬망에 영향을 주는 7가지 단일 위험인자를 방사형 그래프로 나타낸 것이다. 뇌졸중/간질이 가장 높으며, 전해질 불균형과 대사 이상이 그 다음으로 높은 것을 한눈에 확인 할 수 있다. 즉, 단일

위험인자로 보았을 때는 방사형 그림이 작으면서, 전해질 불균형과 대사이상으로 인해서 한쪽으로 치우쳐 있는 것을 확인 할 수 있다.

이전 연구에서는 연령을 제외한 14가지 위험인자를 가지고 베이지안 네트워크를 통해서 7가지 위험인자를 선택하였다 (이제영과 최영진, 2011). 7가지 위험인자로 뇌졸중/간질, 대사이상, 신장수치 이상, 혈액동학적 이상, 기관지 확장제, 전해질 불균형, 항생제가 섬망에 영향을 주는 위험인자로 선택이 되었다. 본 논문은 연령을 제외한 31가지의 위험인자를 가지고 분석을 하였는데, 차이점은 본 논문은 이노제와 음주 스테로이드가 위험인자로 선택되었지만, 이전 연구에서는 항생제와 신장수치 이상, 기관지 확장제가 위험인자로 선택이 되었다

표 4.2는 단일 위험인자를 베이지안 네트워크를 통해 위험인자 간 복합작용을 한 결과를 나타냈다. 단일 위험인자 7가지 중에서 사후확률이 높은 3가지 위험인자 뇌졸중/간질, 전해질불균형, 대사이상을 가지고 다른 위험인자와 복합작용을 통해서 섬망이 나타날 사후확률을 구한 것이다.

표 4.2 주요 위험인자 복합작용 결과

주요위험인자	사후확률	위험인자	사후확률
뇌졸중/간질	0.881	이노제	0.959
		대사이상	0.857
		음주	0.784
전해질불균형	0.352	대사이상	0.798
		스테로이드	0.605
		혈액동학적	0.561
대사이상	0.231	혈액동학적	0.414
		이노제	0.195
		음주	0.128

표 4.2를 보면 뇌졸중/간질만 있을 경우는 88.1%였지만, 다른 위험인자들과 복합작용을 한 결과 각각 95.9%, 85.7%, 78.4%로 사후확률이 나타났다. 이는 뇌졸중/간질 환자가 이노제를 투약한 경우, 대사이상이 있는 경우, 음주를 하는 경우에 대한 사후확률 값이다. 뇌졸중/간질과 복합작용된 이노제와 대사이상, 음주는 단일로만 보았을 때는 7.1%, 23.1% 4.4%로 낮게 나왔지만, 복합작용을 한 결과는 높게 나타난 것을 확인하였다. 전해질 불균형을 다른 위험인자와 복합작용을 한 결과는 단일로만 보았을 때는 35.2%였고, 대사이상은 23.1%, 스테로이드는 0%, 혈액동학적 이상은 18.3%로 나타났다. 전해질 불균형을 가진 사람이 대사이상이 있는 경우와 스테로이드를 투약하는 경우, 혈액동학적 이상이 있는 경우가 각각 79.8%, 60.5%, 56.1%로 높아진 것을 확인하였다. 마지막으로 대사이상을 가진 사람이 혈액동학적 이상, 이노제 투약, 음주를 하는 경우의 확률은 각각 41.4%, 19.5%, 12.8%인 것을 확인하였다. 단일로만 보았을 때는 뇌졸중/간질의 위험인자만 섬망에 걸릴 확률이 50%이상이 되었지만, 위험인자를 복합적으로 가진 경우에는 섬망에 걸릴 확률이 50%넘는 경우가 많은 것을 확인하였다.

그림 4.2는 섬망에 영향을 미치는 7개의 위험인자 중에서 섬망에 영향을 많이 미치는 주요 3개 위험인자인 뇌졸중/간질, 전해질 불균형, 대사이상을 중심으로 다른 위험인자들과의 복합작용 결과를 방사형 그래프를 통해서 나타내었다.

앞에서 본 단일 위험인자 간의 방사형 그림은 모형이 작고, 한쪽으로 치우쳐 있는 것을 확인하였다. 그림 4.2는 그림 4.1의 단일 위험인자 방사형 그림보다 전체적으로 큰 방사형을 이루고 있는 것을 확인할 수 있다. 이는 단일위험인자의 복합작용으로 인해서 섬망에 걸릴 확률이 복합작용이 없는 경우의 확률보다 높은 영향력을 미치고 있는 것을 확인 할 수 있다. 그림4.2를 통해서 위험인자 간 복합작용이 넓게 규명된 것을 시각적으로 쉽게 파악할 수 있었다.

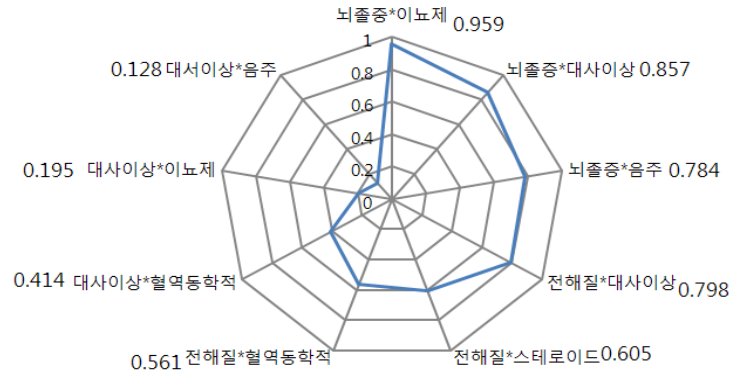


그림 4.2 두 위험인자의 복합작용을 통한 방사형

5. 결론 및 토의

본 논문은 베이저안 네트워크 모형을 적용하여 섬망과 관련된 위험인자를 선별하고 위험인자간 복합 작용을 통해서 방사형 그래프로 규명하였다. 우리는 데이터마이닝 기법인 C5.0, CART, 로지스틱 회귀 모형, 신경망, 베이저안 네트워크를 섬망자료에 적용하였을 때, 가장 모형의 정확도가 높은 베이저안 네트워크를 최종모형으로 선택하였다. 베이저안 네트워크를 최종모형으로 하였을 때, 섬망에 영향을 미치는 위험인자로 뇌졸중/간질, 혈역동학적 이상, 이노제, 전해질 불균형, 대사이상, 음주, 스테로이드 7가지 주요 위험인자가 선택되었다. 단일로 위험인자를 가진 환자의 경우, 섬망이 나타날 확률이 높은 위험인자는 뇌졸중/간질, 전해질불균형, 대사이상을 가진 경우로 나타났다. 위험인자를 복합작용 한 경우에는, 뇌졸중/간질과 이노제 투약한 경우가 섬망 발병률이 가장 높게 나온 것을 확인 할 수 있었다. 대사이상이 있는 경우와 음주를 하는 경우도 뇌졸중/간질과 각각 복합작용을 통해서 단일 보다 상당히 높게 나타난 것을 확인 할 수 있었다. 각각 95.9%, 85.7%, 78.4%로 높게 나타난 것을 확인 할 수 있었다. 즉, 단일 위험인자를 가진 경우보다 복합작용으로 나타날 때 섬망은 훨씬 높게 발병한다는 것을 확인 할 수 있었다. 따라서 우리는 최종모형으로 선택한 베이저안 네트워크를 활용하여서 섬망에 영향을 미치는 31가지 인자 중에서 7가지의 위험인자를 선별하였고, 이들 간에 복합작용 결과를 방사형 그래프로 표현하였다. 또한 베이저안 네트워크 기법을 이용하여서 위험인자가 단일 때 보다는 복합작용을 통한 영향이 섬망에 영향을 많이 미치는 것을 밝힐 수 있었다.

베이저안 네트워크를 통해 밝힌 위험인자 7개를 지닌 환자들을 주의 깊게 살펴야 하며, 특히 7개의 위험인자 중에서 뇌졸중/간질과 전해질 불균형, 대사이상을 가진 환자들을 주의깊게 살펴, 뇌졸중/간질이 있는 사람은 이노제를 투입하지 않도록 주의깊게 살핀다면 사전에 섬망의 발병률이 줄어 들 것이며, 사전에 큰 위험을 막을 수 있을 것이다.

참고문헌

강석복, 박일수, 한준태, 지재훈 (2010). 데이터마이닝을 이용한 위암 예측모형 개발과 활용. <한국데이터정보과학지>, 21, 1253-1261.
 김경현 (2005). < 베이저안 네트워크에 기초한 백혈병 유전자데이터의 분석>, 학사학위논문, 서울대학교, 서울.
 이용원 (2004). <고혈압의 위험요인에 대한 데이터 마이닝 모형 분석-중합건강검진 데이터를 바탕으로>, 박사학위논문, 영남대학교, 경산.

- 이재원, 박미라, 유한나 (2005). <생명과학연구를 위한 통계적 방법>, 자유아카데미, 서울.
- 이제영, 최영진 (2011). 베이지안 네트워크를 활용한 정신장애 질병 섬망(delirium)의 주요 요인 네트워크 규명. <응용통계연구>, **24**, 323-333.
- 허명희, 이용구(2008). <데이터 마이닝 모델링과 사례>, 한나래, 서울.
- Arend, E. and Christensen, M. (2009). Delirium in the intensive care unit : A review. *British Association of Critical Care Nurses, Nursing in Critical Care*, **14**, 145-154.
- Beom, S. K. and Song, K. S. (2009). An empirical study on the roles of attitudes and attitude strength in stimulus-based decision-making. *Journal of the Korean Data & Information Science Society*, **20**, 563-575.
- Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). *Classification and regression tree*, Chapman and Hall, Boca Raton.
- Cole, M. and Primeau, F. (1993). Prognosis of delirium in elderly hospital patients. *Canadian Medical Association Journal*, **149**, 41-46.
- Dubois, M., Strobik, Y., Bergeron, N., Dumont, M. and Dial, S. (2001). Delirium in an intensive care unit : A study of risk factors. *Intensive Care Medical*, **27**, 1297-1304.
- Heckerman, D. (1995). *A tutorial on learning with Bayesian networks*, Technical Report MSR-TR-95-06, Microsoft Research, 1-58.
- Inouye, S. (1994). The dilemma of delirium : Clinical and research controversies regarding diagnosis and evaluation of delirium in hospitalized elderly medical patients. *The American Journal of Medicine*, **97**, 278-288.
- Inouye, S., Schlesinger, M. and Lyndon, T. (1999). Delirium : A symptom of how hospital care is failing older persons and a window to improve quality of hospital care. *The American Journal of Medicine*, **106**, 565-573.
- Jensen, F. (1996). *An introduction to Bayesian networks*, Springer-Verlag, New York.
- Kwak, K., Lee, S. B. and Do, B. S. (2011). Delirium in an emergency department : A study of risk factors. *Journal of the Korean Society of Emergency Medicine*, Submitted.
- Min, D. K. and Hyun, M. S. (2009). Prediction of a winner in PGA tournament using neural network. *Journal of the Korean Data & Information Science Society*, **20**, 1119-1127.
- Tan, P., Steinbach, M. and Kumar, V. (2007). *Introduction to data mining*, Addison Wesley Longman, California.
- Warren, S. (1994). Neural networks and statistical models. *Proceedings of the 19th Annual SAS Users Group International Conference*, 1-13.

The effect investigation of the delirium by Bayesian network and radial graph

Jea-Young Lee¹ · Jae-Young Bae²

¹²Department of Statistics, Yeungnam University

Received 8 August 2011, revised 25 August 2011, accepted 5 September 2011

Abstract

In recent medical analysis, it becomes more important to looking for risk factors related to mental illness. If we find and identify their relevant characteristics of the risk factors, the disease can be prevented in advance. Moreover, the study can be helpful to medical development. These kinds of studies of risk factors for mental illness have mainly been discussed by using the logistic regression model. However in this paper, data mining techniques such as CART, C5.0, logistic, neural networks and Bayesian network were used to search for the risk factors. The Bayesian network of the above data mining methods was selected as most optimal model by applying delirium data. Then, Bayesian network analysis was used to find risk factors and the relationship between the risk factors are identified through a radial graph.

Keywords: Bayesian network, data mining, delirium mental disorder.

¹ Corresponding author: Professor, Department of Statistics, Yeungnam University, Kyungsan 712-749, Korea. Email: jlee@yu.ac.kr

² Graduate, Department of Statistics, Yeungnam University, Kyungsan 712-749, Korea.