

음식점 체인에서 적용 가능한 데이터마이닝

우진형, 김응모
성균관대학교 정보통신공학부
e-mail : woozoo09@paran.com

Datamining for restaurant chain

Jin-Hyung Woo, Ung-Mo Kim
School of information & communication engineering, Sungkyunkwan University

요 약

오늘날 데이터베이스 시스템의 사용은 여러 분야에서 필수적으로 사용되고 있다. 이처럼 다양한 분야에서 삶의 질이 점점 향상됨에 따라 보다 차별적인 서비스를 요구하는 소비자층이 늘고 있다. 이런 변화에 맞춰 본 논문에서는 음식점 체인에서 데이터마이닝 기법을 이용하여 보다 더 효율적이고 유용한 서비스를 제공할 수 있도록 소비자의 행동패턴을 분류하는 방법을 제안한다.

1. 서론

현대는 데이터와 정보의 홍수 시대라고 말해도 과언이 아닐 것이다. 하루에 생성되는 정보와 데이터 량을 감시하는 것조차 불가능해졌고 데이터가 증가할수록 정보는 감소한다는 새로운 역설에 부딪히고 있다. 통계기법이나 간단한 질의만으로도 충분했던 과거와 달리 유용한 정보를 찾아내는 것은 상당한 시간적, 기술적 노력이 필요하다. 현재 대용량의 시스템 안에 수록된 방대한 과거 데이터로부터 뭔가 의미 있는 결과를 끌어내어 좀 더 전략적인 차원에서 데이터를 이용하려는 움직임이 나타나기 시작하였다. 이러한 움직임에 부응해서 등장한 것이 바로 데이터마이닝이다[1].

여기서 제안하는 것은 음식점 체인을 찾는 소비자들의 행동패턴을 데이터웨어하우징을 통해 관리를 하고 데이터마이닝을 통해 소비자가 선호할만한 서비스를 예측함으로써 잠재적 고객을 이끌어낼 뿐만 아니라 기존 고객의 이탈도 막을 수 있다는 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구로 연관규칙 기법에 대해 기술하고, 3장은 시나리오 및 구성을, 그리고 4장에서는 결론 및 향후 발전 방향에 대해 기술한다.

2. 관련연구

2.1 연관규칙 마이닝

연관성 규칙은 데이터마이닝의 초기작업으로서, 자료에 대한 탐색적인 기법으로 유용하게 사용될 수 있는 방법이다. 또한 이것은 마케팅에서 고객이 동시에 구매한 장

바구니를 살펴봄으로써 거래되는 상품들의 관련성을 발견 또는 분석하는 방법으로 장바구니 분석(market basket analysis)이라고도 알려져 있다. 이렇듯 연관성 분석은 연관성 규칙을 통해서 하나의 거래나 사건에 포함되어 있는 둘 이상의 품목들의 상호관련성을 발견하는 것이다. 일반적으로 연관성 분석은 수학과 통계학의 확률과 기대치에 대한 개념을 기반으로 하고 있는데, 이러한 연관성 규칙을 해석하는데 있어 원인과 결과의 직접적인 인과관계로 생각해서는 곤란하고 두 개 또는 그 이상 품목들 사이의 상호의 관련성으로 해석해야 한다[2].

데이터마이닝에서 연관규칙(association rule)은 If A, then B ($A \rightarrow B$)와 같은 형식으로 표현된다. 즉 “상품 A가 구매되었던 경우는 상품 B도 구매된다”라고 해석된다. 본 논문에서는 구체적인 상품들이 언급되므로 보다 실제 상황에 적용이 수월할 것이다[2,4].

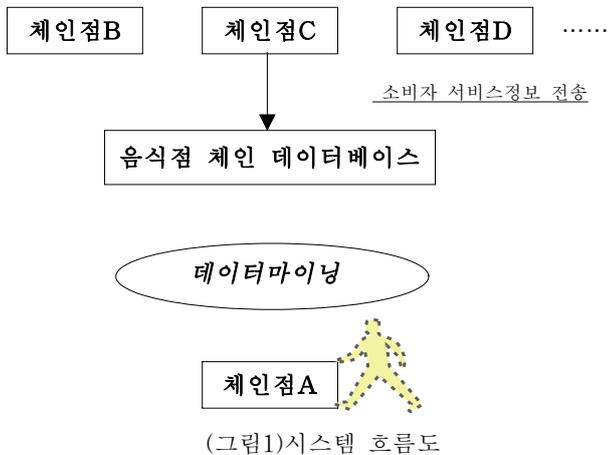
2.2 시차 연관성 규칙

단순히 거래에 포함되는 품목들의 관련성을 살펴보는 문제를 확장해서 동일한 고객이 시간이 지남에 따라서 어떤 소비형태를 보이는가에 대한 분석도 흥미 있는 부분이다. 이러한 분석을 위해서는 거래와 품목으로 정리된 자료로는 부족하고 개개의 고객의 거래내역과 시간순으로 나열된 시계열(time series)자료가 필요하다. 또한 이러한 시간의 연관성을 갖는 규칙들은 원인과 결과의 형태로 해석이 가능해 지므로 그 결과가 더욱 유용할 수 있다. 시간에 관련되어 자료가 정리되므로 실제로 드물게 발생하는 거래의 경우에는 한 달 사이의 거래를 동일한 시점의 거래

로 정의해 자료분석에 이용할 수 있다[3,4].

3. 시나리오 및 구성

음식점 체인을 애용하는 소비자는 연령대별, 성별 등에 따라 선호하는 음식이 다를 것이며 요일별, 혹은 특정 기념일에 따라 부가적인 정보로 분류하면 더욱 정확하게 선호하는 음식을 분류할 수 있을 것이다. 예를 들어 여러 지역의 음식점 체인이 고객이 언제 방문을 하고 무엇을 주로 주문하는지 알기 위하여 고객주문 데이터를 마이닝 할 수 있다. A지역의 음식점 체인에서 녹차를 즐겨 마시는 소비자를 맞이 할 때, 데이터마이닝의 결과를 토대로 음식 주문 전 녹차를 추천함으로써 소비자를 만족시킬 수 있을 뿐더러 판매자의 거래를 증가하는데 사용될 수도 있다.



(그림1)시스템 흐름도

월	주문 내역
6월	오리탕
	유황오리, 부추무침
	콩국수
7월	오리로스구이, 밀전병
	오리백숙, 배추무침
8월	오리야채 무침
	삼계탕, 백반

<표1>소비자1의 고객주문 데이터

어느 소비자의 6월부터 8월까지의 서비스를 데이터웨어하우징을 했을 때 주 메뉴의 재료가 오리고기였다면 음식점 체인의 각 지역에서 고객주문 데이터를 음식점 체인의 데이터베이스를 관리하는 서버로 데이터를 전송한다. 그 후에 그 소비자가 A지역의 체인점을 갔을 때 데이터베이스를 관리하는 서버에서 A체인점으로 해당 소비자의 정보를 전송하면 A체인점에서는 미리 그에 대한 준비를 하여 오리고기를 재료로 하는 메뉴를 추천함과 동시에 곁들여 먹을 수 있는 메뉴도 추가로 추천할 수 있을 것이다. 이는 결과적으로 소비자 입장에서는 양질의 서비스를 받

을 수 있는 것이고 판매자 입장에서 또한 수익을 높일 수 있는 방법인 것이다.

User ID	성별	나이	요리 Key word	디저트 Key word
1	여	35	오리	키위
2	남	22	마늘	바닐라
3	남	24	닭	깻잎
4	남	26	감자	호박
5	여	31	곱창	바닐라
6	여	36	콩	커피

<표2>소비자별 고객주문 데이터

이에 관련한 데이터마이닝은 메인요리 뿐만 아니라 디저트에게까지도 적용될 수 있다. 소비자의 디저트 음식 패턴을 데이터마이닝 하여 메인요리를 먹는 동안 당일의 특별 디저트 메뉴를 준비하여 추천함으로써 음식점 체인에서의 소비자 구매력을 증가시킬 수 있을 것이다.

4. 결론 및 향후 발전 방향

이 논문에서 제안한 고객주문 데이터마이닝에 의한 사전요리메뉴 추천방식은 음식점 체인 이용자가 선호할 만한 음식 정보를 효율적으로 도출하여 보다 높은 만족도를 유지하며 간편한 주문이 가능하도록 하였다. 본 논문에서 처럼 3개월간의 하계 기간 동안 도출된 자료로는 정확성이 떨어질 수 있으니 보다 정확한 데이터마이닝을 위해서는 짧은 기간이 아닌 오랜 기간 동안의 데이터웨어하우징과 데이터마이닝이 필요하다고 생각된다. 또한, 향후 성별, 특정 기념일 등 부가적인 정보를 추가하여 소비자의 구매력을 더욱 끌어낼 수 있는 연구가 진행되어야 할 것이다.

5. 감사의 글

이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임. (No. 2009-0075771).

참고문헌

- [1] 장남식, 홍성완, 장재호 "데이터마이닝" 대청
- [2] 박중수, 유원경, 홍기형 "연관규칙탐사와 그 응용" 정보과학회지
- [3] 강형철, 한상태, 최종후, 김차용, 김은석 "데이터마이닝-방법론 및 활용-" 자유아카데미
- [4] R. Agrawal, T. Imielinski, and A. Swami. "Mining association rules between sets of items in large databases," In Proc. of the ACM SIGMOD