

시계열 인체 센서 데이터의 분류화 기법의 설계와 구현

Xiaoyue Han, 맹보연, 이민수
이화여자대학교 컴퓨터공학과
sowolhan@gmail.com, mngby@ewhain.net, mlee@ewha.ac.kr

Design and implementation of a classification method for time series body sensor data

Xiaoyue Han, Boyeon Maeng, Minsoo Lee
Dept. of Computer Science and Engineering, Ewha Womans University

요 약

무선 통신의 발달과 센서 장비의 소형화로 인하여 다양한 인체 센서들이 개발되고 있으며 이에 따라 이들 인체 센서로부터 생성되는 데이터를 누적하여 분석 및 예측을 해야 할 필요성이 증가하고 있다. 본 연구에서는 누적된 인체 센서 데이터에 대한 분류화 기법을 제안하여 구현하고 성능을 검증하였다. 분류화 기법은 인체 센서 데이터에 잘 적용될 수 있는 지지벡터 기계를 활용하여 구현하였다. 인체 센서 데이터의 대표패턴 정의와 실험을 위한 잡음 생성을 통하여 분류화 정확도를 높일 수 있도록 실험을 설계하였고 다양한 설정 변수에서도 기법을 실험하여 빠르고 정확한 기법을 설계 및 구현하였다.

1. 서론¹

최근 센서와 통신망의 발전으로 언제 어디서든 생체 신호를 수집하여 실시간으로 저장하는 것이 가능하게 되었으나, 병원 내에서 모든 환자의 생체 신호를 모니터링 할 인력과 시간이 부족한 실정이다. 분류화 기법을 통하여 각 사람의 인체 데이터의 특징을 찾아내어 진단이나 치료에 활용할 수 있다.

의료 응용의 구축을 위해 필요한 핵심 기술로서 의료 서버(medical server)에 누적된 생체 및 행위 센서 데이터에 대하여 새로운 패턴을 분석할 수 있는 분류화 기술을 개발하고자 하는 것이 주요한 연구내용이다.

논문의 구성은 다음과 같다. 2 절에서 인체통신망, 시계열 패턴 분석 방법에 대해서 설명한다. 3 절에서는 지지 벡터 기계와 비교 대상인 인공 신경망 혈압 패턴을 정의하고, 지지 벡터 기계를 적용한 시계열 혈압 정보 분류 방법에 대해 설명하고, 4 절에서는 지지 벡터 기계와 인공 신경망에 의한 분류의 성능 비교를 한다. 5 절에서 결론과 향후 연구 과제를 기술한다.

2. 관련연구

2.1 인체 통신망

인체 통신망(Wireless Body Area Network)이란 사람이 착용하는 옷이나 인체에 부착된 여러 디바이스들로 구성된 네트워크를 통해 몸을 중심으로 센서와 구동 기기 간에 데이터 교환을 지원한다. WBAN 을 의료 분야에 이용하여 질병에 대해 사전 검진하고 예방할 수 있으며, 모니터링 시스템을 사용해 만성적인 환자나 노약자에 대해 장기적인 건강 상태를 감지하거나 지속적인 상황을 확인할 수 있다.

2.2 시계열 패턴 분석

시계열이란 시간의 흐름에 따라 일정한 간격으로 관측하여 기록된 자료를 말한다. 분석하고자 하는 조사대상의 자료 값을 일정한 시간간격으로 표시하여, 일정기간 내에 연속된 시점들을 통하여 관측되고, 측정된 일련의 과거 자료들을 일컫는다. 시계열 패턴 분석이란 과거에 발생한 사건은 미래에도 계속적으로 발생할 것이라는 가정을 포함한다. 따라서 시계열 자료의 이용기간이 매우 중요하다. 패턴 분류 시스템의 구성요소는 정보수집, 특징선택, 모델선택, 학습 분류 등으로 이루어진다.

¹ 이 논문은 정부(교육과학기술부)의 재원으로 한국과학재단 중견연구자 사업(No. 2008-0061488)의 지원으로 수행된 연구결과임.

패턴분류의 응용으로는 문자인식, 생체인식과 인간 행동 패턴 분석, 의료영상 분석 및 시스템, 도면인식, 예측 시스템, 보안, 군사 분야에서 사용 된다.

3. 시계열 패턴 분류를 적용한 혈압 정보 분류 방법

3.1 지지벡터 기계

본 논문에서 적용할 지지 벡터 기계는 여백을 최대화하여 일반화 능력을 극대화하는 방법 이다. 패턴에 가장 잘 맞는 여러 지지 벡터들을 찾아내기 위해 다양한 매개 변수의 설정을 한다. 시계열 혈압 패턴은 다양하게 정의 될 수 있기 때문에 선형 분리는 불가능하고 최적의 커널과 매개변수를 휴리스틱한 방법으로 접근해야 한다.

3.2 인공 신경망

인공 신경망은 컴퓨터가 사람의 학습기능을 갖게 하기 위해 고안되었는데 신경망을 구성하는 기본적인 요소로는 신경세포 (neuron), 연결 가중치 (connection weight), 학습규칙 (learning rule) 이 있다.

3.3 시계열 혈압 패턴 분류의 시스템 구성

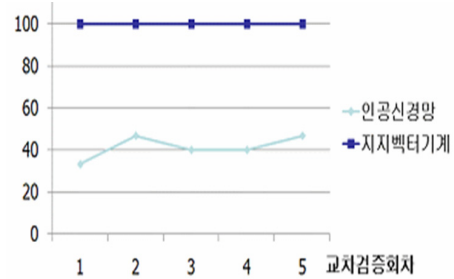
사용자의 혈압 정보는 인체 통신망을 통해 수집되고 혈압정보가 데이터베이스에 누적되고, 저장된 정보의 정규화 과정을 통해 의미 있는 정보를 추출한다. 추출된 정보는 본 논문에서 제안하는 지지 벡터 기계의 모델을 구축하기 위해 학습을 한다. 학습이 끝난 후, 모델을 통해 혈압 패턴이 분류 된다.

3.4 시계열 혈압 패턴 정의

정상인의 혈압 데이터에서 15 개의 패턴을 정의했다. 실제로 데이터가 분석할 수 있을 만큼 쌓여 있지 않기 때문에, 실제 데이터와 비슷하게 하기 위하여 각 패턴당 10 개의 잡음을 추가하고 150 개의 패턴을 사용하였다. 잡음은 알파 값을 더하고 빼는 방법으로 추가하였다. (알파 값:1, 5, 10, 15, 17, 20).

4. 분류화 기법의 성능 비교

실험을 통하여 지지 벡터 기계에서 분류율과 수행 속도 면에서 가장 좋은 Radial Basis 함수를 커널함수로 사용하고 그 매개변수를 0.01 로 구성한다. 인공 신경망에서는 최적의 구성을 갖는 은닉층과 그에 따른 노드 수를 기준으로 하여 시계열 혈압 데이터 분류화에 대한 정확도와 수행속도의 비교를 한다. 이 때, 성능 평가 비교를 위해서 10-fold의 교차 검증을 한다.



(그림 1) $\alpha=15$ 일때 정확도 비교

(그림 1)에서와 같이 지지 벡터 기계는 항상 같은 정확도를 보이지만 인공신경망은 분류를 할 때마다 다른 정확도를 보였다. 지지 벡터 기계는 인공 신경망보다 50.67% 이상의 정확도를 보였다. 다음은 지지벡터 기계는 인공 신경망에 비해 약 90 배 이상의 의 수행 속도를 갖는다.

5. 결론 및 향후 연구

본 논문에서는 지지 벡터 기계를 적용하여 시계열 혈압 패턴을 분류하는 방법을 제안한다. 지지 벡터 기계는 다양한 매개변수의 실험을 통하여 최적화하여야 하나 학습 시간 및 분류화 시간이 인공신경망에 비하여 우수하다. 향후 연구내용으로는 다양한 인체 센서 데이터에 대한 적용을 실험하고 성능비교를 하여야 할 것이다.

참고문헌

- [1] H. Ren, M. Meng, X. Chen, H. Sun, B. Fan, and Y. Chan, "System Architecture of a Body Area Network and Its Web Service Based Data Publishing", Lecture Notes in Computer Science, Springer Berlin, vol. 3842, pp. 1611-3349, 2006.
- [2] M. Lan, A. Nahapetian, A. Vahdatpour, L. Au, W. Kaiser, and M. Sarrafzadeh, "SmartFall: an automatic fall detection system based on subsequence matching for the SmartCane", In Proc of. ICBAN-04, Los Angeles, California, USA, article no. 8, 2009.
- [3] A. D. Jurik, J. F. Bolus, A. C. Weaver, B. H. Calhoun, and T. N. Blalock, "Mobile health monitoring through biotelemetry", In Proc of. ICBAN-04, Los Angeles, California, USA, article no. 12, 2009.
- [4] J. G. Ko, T. Gao, and A. Terzis, "Empirical Study of a Medical Sensor Application in an Urban Emergency Department", In Proc of. ICBAN-04, Los Angeles, California, USA, 2009.
- [5] J. D. Hamilton, "Time Series Analysis", Princeton University Press, 1994.
- [6] L. Cao, "Support vector machines experts for time series forecasting", Neurocomputing, vol. 51, pp. 321-339, April 2003.