

영유아 체온 데이터 기반 빅데이터 분석 및 학습을 위한 데이터 수집 시스템 구현

이현섭 · 허경용*

동의대학교

Implementation of a data collection system for big data analysis and learning based on infant body temperature data

Hyoun-Sup Lee · Gyeongyong Heo*

DongEui University

E-mail : lhskmj@deu.ac.kr / jgycap@deu.ac.kr

요 약

최근 다양한 분야에서 인공지능 시스템이 활용되고 있다. 인공지능의 결정 알고리즘의 정확도는 학습량과 학습데이터의 정확도에 기인한다. 학습량의 경우 인공지능 성능에 결정적인 영향을 미치기 때문에 많은 양의 데이터가 필요하다. 학습데이터의 정확도는 여러 정제 단계를 거치면서 보정할 수 있으나 분석 이외의 자원 소모를 추가로 가져온다.

본 논문에서는 영유아의 체온 데이터를 기반으로 향후 나타날 수 있는 병증 및 유아의 상태 변화를 분석하는 시스템 구축을 위한 데이터 수집 시스템에 대하여 제안한다. 제안된 시스템은 기존 빅데이터 분석 및 학습 데이터 구축에서 서버 시스템의 자원 소모를 최소화할 수 있을 것으로 사료 된다.

ABSTRACT

Recently, artificial intelligence systems are being used in various fields. The accuracy of the decision algorithm of artificial intelligence is greatly affected by the amount of learning and the accuracy of the learning data. In the case of the amount of learning, a large amount of data is required because it has a decisive effect on the performance of AI.

In this paper, we propose a data collection system for constructing a system that analyzes future conditions and changes in infants' conditions based on the body temperature data of infants and toddlers. The proposed system is a system that collects and transmits data, and it is believed that it can minimize the resource consumption of the server system in existing big data analysis and training data construction.

키워드

AI, Data collection, Data cleansing, Mobile App

I. 서 론

인공지능은 다양한 분야에 적용되어 사용되고 있다[1][2]. 인공지능의 성능 여부는 결정 알고리즘의 정확도를 기반으로 판단된다. 결정 알고리즘은 빅데이터 분석을 기반으로 다양한 학습데이터를 생성하여 이를 인공지능 학습에 적용한다. 여기서 빅데이터 분석에 들어가는 시스템의 성능은 일반

서버급 시스템의 모든 자원을 활용하더라도 상당한 시간이 소요된다. 본 논문에서는 이러한 자원 소모를 최소화하기 위해서 분석 과정의 일부를 수행하여 처리하는 수집 시스템에 대하여 제안한다.

II. 관련 연구

게임 인공지능에서 자원 최적화를 위한 유닛의 부하분산 시스템에 관한 연구[3]에서는 복잡한 AI

* corresponding author

를 동작하기 위해서는 서버에 CPU 자원을 효율적으로 사용하는 기법에 대하여 제안하였다.

위와 같이 AI의 처리 성능을 효율적으로 높이는 기법은 다양하게 연구되고 있다. 그렇지만 raw 데이터의 처리와 학습 시점에서의 자원 효율화에 관한 연구는 많지 않다. 많은 시스템이 IOT 기반으로 동작하고 수많은 센서가 실시간으로 많은 양의 데이터를 제공하는 시스템들이 점차 확대되고 있는 상황에서 raw 데이터의 효율적인 처리는 인공지능에 있어서 고려되어야 하는 필수 항목이다.

III. 데이터 수집 모듈 설계

그림 1은 ㈜이즈프로브의 모바일 애플리케이션을 통해 체온계와 연동하여 사용자의 체온의 변화량을 서버로 전달해주는 시스템의 설계 화면이다.

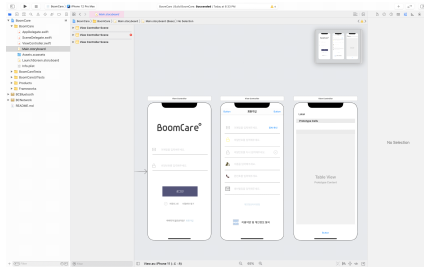


그림 1. 붐케어 시스템의 IOS 애플리케이션

사용자의 측정 체온 정보를 실시간으로 서버로 전달하고 이를 시계열 기반으로 저장하고 분석한다. 분석은 raw 데이터 정제, 전처리, 분석의 세 단계로 나누어지며 온도 변화의 실시간 구간에 따라 분석 레벨은 다양하게 구성되어 있다. 현재 실시간 사용자의 수가 1000여 명 안팎으로 분석 과정에 큰 자원 소모가 없다.



그림 2. 정제 처리 이후 전송된 체온 데이터

그러나 코로나 발병 이후 체온계의 판매량이 급증하였고 이와 함께 애플리케이션을 사용하는 사용자의 수가 크게 증가하여 분석 과정에 시간이 많이 소요되는 특징을 보이고 있다. 이를 위해 애플리케이션에서 결측치와 이상치를 전처리 정제하는 모듈을 그림 2와 같이 구현 중이다.

체온계의 전송 모듈 오류로 발생한 결측치를 제거하고 일반적인 체온 30~40 사이의 범주 데이터의 경우 전송은 하되 이상치 데이터로 전송 과정

에서 생략하거나 플래그를 삽입하여 전송한다.

IV. 결 론

본 논문에서는 인공지능 시스템에서 데이터수집 모듈 기반의 raw 데이터 전처리 시스템에 대하여 제안하였다. 제안하는 시스템의 경우 인공지능의 데이터 분석 과정에 있는 정제과정을 제거하거나 빠르게 처리할 수 있도록 1차 raw 데이터 처리를 수집 모듈에서 수행한다. 이처럼 수집 시스템에서 1차로 raw 데이터를 처리할 경우 분석 및 학습을 수행하는 서버의 자원 부하를 줄이고 빠른 분석 결과를 얻을 수 있을 것으로 사료 된다.

IOT 기반의 다양한 센서 환경에서도 게이트웨이 기반 및 Edge 컴퓨팅 기반의 선 전처리 과정을 도입할 경우 인공지능 분석 과정에서 효율적인 자원 관리가 가능할 것으로 판단 된다.

Acknowledgement

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 Grand ICT연구센터지원사업의 연구결과로 수행되었음” (IITP-2020-0-01791)

References

- [1] S.J.Jun, “Development of Artificial Intelligence Education Program based on Experiential Learning for Liberal Art Education,” The Journal of Korean Association of Computer Education, Vol.24 No2, pp. 63 – 73, 2021.
- [2] S.J.No, J.W.Kim, Y.K.You, J.S.Lee, S.H.Shin, K.B. Lee, “Current Status and Prospect of Artificial Intelligence Utilization in Robot Industry,” The Journal of The Korean Institute of Communication Sciences, Vol.38, No.4, pp. 39-44, 2021
- [3] W.B.Lee, C.Go, “Load Distributed System of Units For Game AI's Resource Optimization,” Journal of the Korean Society for Industrial and Applied Mathematics IT Series, Vol. 8, No. 2, pp. 33-45, 2004