

학습 상담 내용의 자연어 처리를 위한 오픈 데이터 현황 분석

김유두

한국폴리텍대학 융합기술교육원

Analyze the Open data for Natural Language Processing of Learning Counseling

Yu-Doo Kim

Convergence Technology Campus of Korea Polytechnics

E-mail : yudookim@kopo.ac.kr

요 약

4차산업의 융합시대를 맞이하여 단순한 학습이 아닌 다양한 학문을 학습하고 융합하여 적용하기 위해서는 주입식 수업보다는 자기주도방식의 학습방법이 중요해지고 있다. 따라서 다양한 교육 기관에서는 자기주도적인 학습 방법의 개발에 많은 노력을 하고 있다. 자기주도적인 학습이 효과적으로 수행되기 위해서는 교수자는 학생의 학업에 직접 관여하기 보다는 학업의 전체적인 과정을 관리하는 것이 더 중요하다. 이에 학습 상담은 자기주도적학습을 효과적으로 수행하는데 중요한 방법이 된다. 이에 본 논문에서는 학습 상담 내용을 자연어 처리를 통해 다양한 응용이 가능하도록 이를 구현할 수 있는 자연어 처리를 위한 오픈 데이터 현황에 대한 분석을 수행 하였다.

ABSTRACT

In the 4th generation industry, self-directed learning is very important than Injection learning. Therefore many educational institutions has developed method of self-directed learning. In order for self-directed learning to be effective, it is more important for faculty to manage the overall process of learning rather than being directly involved in the student's academic work. Therefore, learning counseling is an important way to effectively carry out self-directed learning. In this paper, we analyze the status of open data for natural language processing that can implement the learning consultation contents so that various applications can be done through natural language processing.

키워드

AI, Open data, big data, data set

1. 서 론

자기주도적 학습이 올바르게 수행되기 위해서는 학습자에게 너무 많은 방법을 제시하거나 공통적인 방법을 제안하기 보다는 학습자 각각의 특성을 파악하고 그에 맞게 교육이 이루어져야 한다.

이에 맞게 공동연구를 수행하고 있는 Y사에서는 모든 학생들에게 지속적인 상담을 통해 학생 개개인에 맞는 공부 방법과 스케줄을 제시하여 공무원 시험 합격률을 높이는 방법으로 자기주도적

학습을 운영하고 있다. 하지만 이러한 방법이 모두 상담하는 지도교사의 수동적이고 개인적인 판단에 의해 이루어지기 때문에 이 부분을 인공지능 기술을 활용한다면 더욱 효과적인 상담이 진행될 수 있다.

이를 위해 본 연구에서는 상담내용을 녹음하거나 문자로 작성하여 저장된 것을 활용하여 학습 상담내용을 분석하기 위한 자연어 처리 오픈 기술 및 데이터를 알아보고 적용할 수 있는 방안을 연구하고자 한다.

II. 학습 상담 내용의 분석

Y사의 학습 상담 내용을 활용하기 위해서 우선, 학습 상담을 하는 내용을 다음과 같이 카테고리화 구분할 수 있다.

표 1. Y사의 학습상담 내용

상담구분	상담내용
학습과목	우수/취약 과목 분석
학습태도	학습 태도 분석
일상상담	일상적인 상담
학습일과	학습 일과 패턴 분석

표에서 보는 바와 같이, Y사의 학습 상담내용은 크게 네 가지로 구분하고 있다. 이와 같이 각 상담 내용의 카테고리를 분류해 놓고, 이를 자연어처리로 분석하여 학생들에게 가장 알맞은 자기주도적 학습 방법을 제안할 수 있게 된다.

III. 자연어 처리를 위한 오픈 데이터

작성된 상담내용을 토대로 분석을 하기 위해서는 자연어를 처리하여 주는 기술과 이를분석하기 위한 데이터가 필요하다. 기존의 학습 상담 내용을 모두 포함하여 학습을 수행하기도 하겠지만, 더 많은 데이터가 있어야 자연어처리가 더욱 잘 동작할 수 있기 때문에 다음과 같은 오픈 데이터와 API를 활용할 수 있다

다음은 정부에서 운영하는 AI오픈이노베이션허브에서 제공하는 데이터의 형태를 보여준다.

표 2. AI오픈이노베이션허브의 서비스

서비스명	내용
언어분석	형태소 분석, 개체명 인식, 동음이의어/다의어 분석, 의존 구문분석, 의미역 인식
어휘관계	어휘에 대한 정보를 검색하고, 어휘들 간의 의미적 연관성을 분석
음성인식	음성인식과 발음 평가
대화처리	대상 도메인 의 목적에 맞추어 대화가 가능하도록 하기 위해 대화 지식이나 데이터를 구축
시각지능	객체 검출 (Object Detection)
질의분석	사용자의 자연어 질문의 의도를 파악하고 사용자가 요구하는 정답을 찾을 수 있도록 정보를 분석

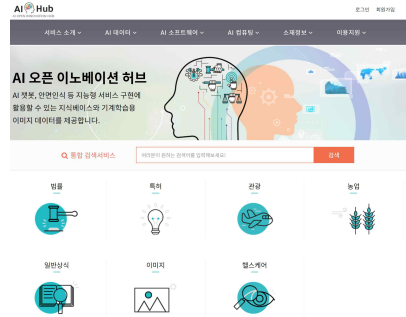


그림 1. AI오픈이노베이션허브 카테고리

지원하는 오픈 데이터셋을 통해 학습을 수행하고, Y사의 기존 상담내용을 추가로 학습한다면 더욱 정확한 분석이 가능한 형태로 머신러닝이 수행될 준비가 될 수 있다. 그 후에 자체적인 AI기능을 개발하지 않더라도 자연어를 처리하여 주는 다양한 오픈 API를 사용하여 분석한다면 학습상담 내용 작성만 사람이 수행하고, 그에대한 학생의 자기주도적 학습을 지원하는 다양한 제안은 인공지능이 스스로 수행할 수 있는 플랫폼을 구축할 수 있다. 이를 통해 조금 더 과학적인 자기주도적학습을 위한 도구로 활용이 가능할 수 있게 된다.

IV. 결 론

자기주도적학습은 학습자 개인의 특성을 반영하여 그에 맞는 학습 방법을 제공해주고 각자의 특성에 맞게 최대한의 학습 효과가 나타날 수 있도록 하는 것이 가장 중요하다. 이를 위해 현재 수동으로 상담 지도교사에 의해 진행되고 있는 학습 제안을 인공지능을 활용하기 위해서 필요한 오픈 데이터와 관련 기술을 알아보고 이를 활용하여 개발할 수 있는 기반을 갖추기 위한 방법을 알아보았다. 이를 통해 자기주도적학습이 더욱 효과적으로 수행되는데 많은 도움이 될 수 있을 것이다.

Acknowledgement

본 연구는 2019년도 중소벤처기업부의 기술개발 사업 지원에 의한 연구임 [S2655756]

References

[1] 7 Amazing Open Source NLP Tools to Try With Notebooks in 2019 : <https://medium.com/microsoftazure/7-amazing-open-source-nlp-tools-to-try-with-notebooks-in-2019-c9eec058d9f1>
 [2] AI OPEN INNOVATION HUB : <http://http://www.aihub.or.kr/>.