**ChatGPT와 딥 러닝을 활용한 지리 정보 처리: 개체명 인식과 개체명 연결 성능 비교 연구**

**"Comparative Study of Geographical Information Processing with ChatGPT and Deep Learning: Performance Comparison of Named Entity Recognition and Named Entity Linking"**

최대웅

Dae Woong Choi

서울대학교 스마트도시공학과 석사과정 ([cdwoong828@snu.ac.kr](mailto:cdwoong828@snu.ac.kr))

**요약**

본 논문은 ChatGPT와 딥 러닝 기술을 이용하여 지리 정보 처리 분야에서 개체명 인식과 개체명 연결 기술의 성능을 비교하는 연구를 수행하였다. 본 연구에서는 지리 정보 처리의 중요성과 그 응용 분야를 강조하며, 개체명 인식과 개체명 연결 기술의 성능이 지리 정보 처리 분야에서 매우 중요하다는 것을 목적으로 한다. 실험에서는 BERT와 같은 딥 러닝 기반의 모델을 사용하여 개체명 인식을 수행하였으며, fuzzy string match와 같은 전통적인 방법을 사용하여 개체명 연결을 수행하였다. ChatGPT와 같은 데이터를 사용하여 결과를 비교 분석하였으며, 개체명 인식과 개체명 연결 모두에서 딥 러닝 기반 모델이 기존의 전통적인 방법보다 더 우수한 성능을 보였다또한, 개체명 인식과 개체명 연결 모델 간의 성능 차이를 분석하였으며, 개체명 인식 모델의 성능이 개체명 연결 모델의 성능에 직접적인 영향을 미치는 것을 확인하였다. 이러한 결과들은 지리 정보 처리 분야에서 딥 러닝을 활용한 개체명 인식과 개체명 연결 기술의 중요성을 더욱 부각시키는 의의를 가지고 있다.

1. **서론**

본 연구는 지리 정보 처리 분야에서의 자연어 처리 기술을 활용하여 개체명 인식과 개체명 연결 성능을 비교하는 것을 목적으로 한다. 지리 정보 처리 분야에서는 지명, 기관명, 장소명 등 다양한 개체명 정보를 다루어야 한다. 따라서, 정확한 개체명 인식과 연결이 필요하다. 그러나 이를 위해서는 다양한 언어적, 문화적, 지리적 요소들을 고려해야 한다. 이러한 문제점으로 인해, 지리 정보 처리 분야에서는 기존의 자연어 처리 기술의 성능이 한계에 도달하였다.

개체명 인식과 개체명 연결은 자연어 처리 분야에서 중요한 문제이다. 개체명 인식은 텍스트에서 개체명을 식별하는 작업으로, 예를 들어 "서울특별시 강남구 역삼동"이라는 문장에서 "서울특별시", "강남구", "역삼동"이 개체명이다. 개체명 연결은 서로 다른 문장에서 같은 개체명을 식별하여 연결하는 작업으로, 예를 들어 "역삼동"이라는 개체명이 포함된 문장이 여러 개 있다면, 이들 문장에서 같은 "역삼동" 개체명을 식별하여 연결해야 한다. 개체명 인식과 연결은 다양한 자연어 처리 분야에서 매우 중요한 작업으로 활용되고 있다.

ChatGPT는 OpenAI에서 최근에 공개한 대형 언어 모델로, 매우 높은 성능을 보여준다. 딥 러닝은 인공신경망을 이용하여 자연어 처리 분야에서 다양한 작업을 수행하는 기술로, 성능이 뛰어나기 때문에 최근에는 많이 활용되고 있다. 이러한 ChatGPT와 딥 러닝 기술은 지리 정보 처리 분야에서의 자연어 처리 기술에도 적용할 수 있으며, 개체명 인식과 개체명 연결 분야에서의 성능 개선에 많은 기여를 할 수 있다.

지리 정보 처리는 우리 일상생활에서 매우 중요한 역할을 하고 있다. 이를 기반으로 하는 위치 기반 서비스는 급속하게 발전하면서 많은 기업들이 경쟁하고 있다. 이러한 서비스에서 개체명 인식과 연결은 지리 정보 처리의 정확도와 유용성에 직결되기 때문에 매우 중요한 문제이다. 그러나 기존의 자연어 처리 기술은 다양한 언어적, 문화적, 지리적 요소들을 고려하지 못해 성능이 한계에 도달하였다. 이에 딥 러닝과 ChatGPT와 같은 대형 언어 모델이 등장하면서 자연어 처리 분야에서 많은 성과를 이루고 있다.

본 연구에서는 ChatGPT와 딥 러닝 기술을 활용하여 개체명 인식과 개체명 연결 성능을 비교하여, 이를 통해 지리 정보 처리 분야에서의 성능 개선 방안을 모색하고자 한다. 특히, 개체명 연결에서는 fuzzy string match와 같은 전통적인 방법과 ChatGPT를 이용한 방법을 비교해보고자 한다. 이를 통해 자연어 처리 분야에서의 딥 러닝과 대형 언어 모델의 성능과 활용 가능성을 지리 정보 처리 분야에서도 검증하고자 한다.

1. **본론**

본 연구에서는 ChatGPT와 RNN, LSTM, Attention 기반의 딥 러닝 방법을 사용하여 개체명 인식과 개체명 연결 성능을 비교합니다. 이를 위해 오픈 소스 지리 정보 데이터 세트와 지리 정보 처리 분야에서 일반적으로 사용되는 평가 척도를 사용합니다. ChatGPT는 최근 대규모 언어 모델의 개발로 인해 주목받고 있는 자연어 처리 기술 중 하나이며, 자연어 처리 분야에서 상위 성능을 보여준 모델입니다.

따라서 본 연구는 ChatGPT를 포함한 딥 러닝 방법과 기존 머신 러닝 방법의 개체명 인식과 개체명 연결 성능을 비교하고, ChatGPT를 이용한 모델이 더 나은 성능을 보인다는 것을 보여줌으로써, 자연어 처리 기술이 지리 정보 처리 분야에서의 활용 가능성을 제시하는 것을 목적으로 합니다.