



신경/혈관 도메인 SRA 연구 추진안

(주) 커넥티드인 부설연구소
권찬우 / 연구소장



Connected In Inc. R&D Center since
2022

- 개요 Outline
- 현재 상황 Current situation
- 장점 Advantage
- 단점 및 해결방안 Disadvantage and solution
- 데이터 분석 방안 The way to analyze data

- 신경/혈관에 관한 병리학적 접근, 연구방법론, 생리학적 접근 등 여러 연구자료에 대한 체계적 문헌고찰의 자동화를 연구 및 개발

Research & development of SRA for various research data such as pathological approach, research methodology, and physiological approach related to nerves/vasculature

- 기존 SRA 와의 차별성이 중요

The differentiations are the most important

- 신경/혈관 도메인에 특화된 데이터 분석

Analyzing the specific data for nerves/vasculature domain

- 논문간 관계를 통해 다시각적 분석이 가능한 결과 도출

Draw results to enable analysis from various perspectives through the relationship between papers

- 체계적 문헌고찰 자동화(SRA)의 차별성 필요

Need differentiations with other SRA

- 특정 도메인 의존적인 SRA 개발이 차별화에 유리

Developing specific domain dependent SRA is advantageous for differentiation

- 신경/혈관 연구자료 확보 (약 25만 건, 단, 중복처리 하지 않음)

We already got the research data (Almost 250K, but can be duplicated)

- 이미 많은 수의 논문 확보 (약 25만 건)

We already got lots of paper (almost 250K)

- 개발/테스팅 시 시간을 크게 단축

Reduce the time for development/testing

- 정밀의학 교수의 도움을 받을 수 있음

We can get help from the professor of precision medicine

- 개발 개요/과정/결과 등 여러 요소에 자문을 구할 수 있음

We can get advice such as outline/process/result, etc.

- 특정 데이터 분석으로 다양한 효과 기대

Expect various effects from analysis of specific data

- 단어 유사성 뿐만 아니라 해당 도메인의 다양한 데이터 분석

Analyze various data of that domain not only the word similarity

- 데이터의 관계 분석을 통해 논문간 관계 추론

Inference the relationships between papers from analysis the relationships between data

- 다시각적 분석이 가능하도록 결과 도출

Draw results to enable analysis from various perspectives

- 문헌고찰 시 다양한 사고방식이 가능할 것으로 기대

Expect various ways of thinking will be possible when reviewing

- 도메인 지식 부족

Lack of domain knowledge

- ✓ 기존 NLP 알고리즘을 통해 전문 용어 분석

Analyze the medical terms using existing NLP algorithms

- ✓ 전문 용어 분석을 통해 실제 분류 결과의 정확성 파악

Determine the accuracy of the clustering results through the analysis of the medical terms

- ✓ 단어 유사성 □ 전문 지식 없이 통계로 분석 가능

Word similarity □ Can be able to analyze from statistics without expertise

- 이미 개발된 SRA와의 차별성

The differentiations from the existing SRA

- ✓ 기존 SRA는 특정 의학도메인이 아님

Existing SRAs are not for specific medical domain

- ✓ 신경/혈관에 관한 전문적 SRA

Our SRA is special for nerves/vasculature

- ✓ 다시각적 분석이 가능한 결과 도출

Draw results to enable analysis from various perspectives

- 이미지/데이터 분석이 힘들

Too hard to analyze the image/some data

- ✓ 이미지의 경우 매핑되는 단어 존재 – ex) 망막 혈관 사진이 나오는 논문은 망막에 관한 의학용어 출현

All image data have mapping word – ex) If the paper has retina vascular image, also has medical term related to retina, too

- ✓ 의학용어, 분자 화학식 등의 데이터와 함께 NLP 알고리즘 응용 가능

Can be applicated using NLP algorithms with medical terms, chemical formula, etc.

- 판단 결과 기준 모호, 결과물 효과성 판단 모호

Ambiguous criteria for judgment results, ambiguous judgment of outcome effectiveness

- ✓ 정밀의학 교수의 자문을 구할 수 있음

We can get advices from professor of precision medicine

- 세포, 단백질 관련 화학식 및 분자 이름

Cells, chemical formula related to protein, and molecule name

- 세포 관계도 구축 □ 세포 구성요소, 반응 물질 등 데이터베이스 구축 및 관계 분석 (유사, 연결, 협력 등)

Build relationship map of cells □ Setup the database about cell components, metabolism, etc. and analyze relationships (similarity, connection, co-working, etc.)

- 단백질 화학식 그래프 구축 □ 화학식의 벡터화 및 관계 분석 (포함, 유사, 동시 출현 등)

Build the graph from chemical formula of protein □ Vectorization of chemical formula and analyze the relationships (include, similarity, co-occurred, etc.)

- 망막 등 혈관 이미지, 신경세포 이미지

Vascular image such as retina, etc. and nerve cell image

- 단어 대체 가능

Can use the medical terms instead of image

- 혈관 분류, 패턴 분석 등 Challenge project 확장 가능

Can make challenge project such as classification the vascular, analyzing the pattern of vascular or nerve cells

- 질환명, 약품명, 수술방법과 같은 의학용어

The medical terms such as diseases, medicine, surgery, etc.

- 의학 용어 사이의 관계 분석 (CDM 활용)

Analyze the relationship between the terms (using CDM)