



이기창, Gichang Lee
ratsgo@naver.com
ratsgo.github.io
010-9692-8870

주소
경기 용인 동천동 899
분당수지유타워
B동 417호

이기창

관심 분야 임베딩, 문장생성 등에 관심이 많습니다.

학력

2016 - 2018, 고려대학교

공학석사(산업경영공학)

2003 - 2010, 서울대학교

문학사(국어국문학), 정치학사(정치외교학)

경력

네이버 서치엔클로바 (2018.2 - 현재)

- 담당 업무 : 대화모델 연구&개발

매일경제신문 (2009.8 - 2016.6)

- 직위 및 담당 업무 : 편집국 기자, 취재 및 기사 작성
- 담당부서 : 사회부, 경제부, 정치부, 문화부, 산업부
- 주요 기사
 - 천안함 생존장병들의 부치지 못한 편지 [link]
 - G20 서울선언문 입수, '햇머니 규제' 합의안 담는다 [link]
 - 한나라당 2조 공약안 입수, 모든 가구에 23만원 양육수당 [link]
 - 박근혜 집권시 창조경제론 전면 내세운다 [link]
 - 교황 시복식 현장 취재기, 소란 없는 '엄숙한 축복' [link]
 - 삼성 갤럭시 새 사령탑 고동진 깜짝 발탁 [link]

논문 및 저서

이기창 (2019). 한국어 임베딩 (Sentence Embeddings Using Korean Corpora). 에이콘출판사. [link]

Gichang Lee, Jaeyun Jeong, Seungwan Seo, CzungYeob Kim, Pilsung Kang. (2018). Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network. Knowledge-Based Systems. Volume 152, 15 July 2018, 70-82. (SCI) [link]

서덕성, 모경현, 박재선, 이기창, 강필성. (2017). 워드임베딩과 그래프 기반 준지도학습을 통한 한국어 어휘 감성 점수 산출. 대한산업공학회지 (JKIIE), 43(5), 238-248. [link]

조수현, 김보섭, 박민식, 이기창, 강필성. (2017). 여행 사이트 리뷰를 활용한 관광지 만족도 요인 추출 및 평가, 대한산업공학회지 (JKIIE), 43(1), 62-71. [link]

발표

이기창, 정재윤, 서승완, 김창엽, 강필성. (2017). 합성곱 신경망을 사용한 약지도학습 기반의 감성분류 및 단어 어텐션. 대한산업공학회 추계학술대회, KAIST, 대전, 11월 4일.

박민식, 이기창, 김보섭, 김창엽, 강필성. (2017). 워드 임베딩을 활용한 휴대폰 리뷰에서의 핵심 대상 기능 추출. 대한산업공학회 춘계학술대회, 여수 EXPO 컨벤션센터, 여수, 4월 26일-4월 29일.

이기창, 강필성. (2017). Graph-based Representation을 활용한 뉴스 중요도 산출. 대한산업공학회 춘계학술대회, 여수 EXPO 컨벤션센터, 여수, 4월 26일-4월 29일.



이기창, Gichang Lee
 ratsgo@naver.com
 ratsgo.github.io
 010-9692-8870

주소
 경기 용인 동천동 899
 분당수지유타워
 B동 417호

연구

정형/비정형 데이터 기반의 경제 Knowledge Mining 기술 연구

(2017.5 - 2018.3) (주)엔씨소프트 지원 과제

- 학생 연구원으로 참여
- 기여 내용 : 조간신문 1-5면 제목을 꼭지점, 중복단어를 간선으로하는 무방향 그래프 구축 후 Edge Betweenness Score를 기준으로 그래프 컷. 이와 별도로 기사 별로 $\log(\text{기사길이})/\text{면수}$ 를 계산한 뒤 이 점수가 높은 부분그래프 10개를 선택. 부그래프에서 고유벡터 중심성이 가장 높은 제목을 해당 날짜의 중요 이벤트로 정의.

분포와 의미 정보를 동시에 반영한 한국어 단어 임베딩

(2017.4 - 2017.10) 개인 연구

- 국립국어원이 만든 세종전자사전의 한국어 명사 온톨로지(2만5000여개 태깅)를 그래프로 표현한 뒤, Grover&Leskovec(2016)이 제안한 Random Walk를 뽑고, 뉴스 말뭉치와 함께 Word2Vec Skip-Gram 방식으로 임베딩을 수행. 명사 온톨로지의 지식 표상과 뉴스 말뭉치(55만여건)를 동시에 임베딩해 분포와 의미 정보를 동시에 반영.

기술 블로그 구축 및 운영

(2017.3 - 2018.2) 개인 연구

- 국어학, 자연언어처리 등 주제를 다루는 GitHub 블로그 구축. 2018년 6월 현재 게시물 216건, 월 평균 방문자 3만2000여명.

비정형의 빅데이터를 활용하여 경쟁환경에서 사업자의 신제품 포지셔닝과 다기간 제품 포트폴리오 선정에 도움을 줄 수 있는 전사적 의사결정 시스템의 개발

(2016.9 - 2016.12) 한국연구재단 지원 과제

- 학생 연구원으로 참여
- 기여 내용 : 소비자가 중시하는 제품 속성을 분석. 우선 뽀뿌 등 5개 사이트에서 29만8206개 리뷰를 수집해 Document Term Matrix(DTM) 구축. Word2Vec으로 임베딩한 단어벡터에 가우시안 커널을 적용하여 가중치 행렬을 계산. 배터리, 카메라, 디자인, 화면이라는 단어에 해당하는 4개 행벡터를 뽑고 이를 DTM과 내적해 스코어를 각각 산출. 스코어 상위 문장을 랜덤 추출해 100명을 대상으로 교내 설문조사를 한 결과 단순정확도 기준 73.1%의 성능을 보임.

그래프 기반 뉴스 중요도 산출

(2016.7 - 2016.10) 개인 연구

- 다수 보도된 유사한 제목의 기사를 중요한 기사로 간주. 제목을 꼭지점, 중복 단어를 간선으로 하는 그래프 구축 후 간선중심성과 고유벡터 중심성을 곱해 중요도 산출. 중요도 점수 상위 3% 기사 가운데 1면은 16.97%로 전체 기사 가운데 1면 비율(4.1%)보다 4.13배. 파급력이 큰 뉴스를 보편적으로 전달해야 한다는 가치가 특종 보도 원칙을 앞서는 국내 언론 환경 때문으로 풀이.

기술

프로그래밍

- Python
- Scala
- R

작문

- 취재
- 자료분석
- 원고작성