

# 김명섭

---

## 연락처

- Email: [Msub0310@gmail.com](mailto:Msub0310@gmail.com)
  - Phone: 010-5597-6777
  - Github: <https://github.com/msub0310>
- 

## 관심 분야

- 소셜 미디어, 특히 등의 오픈 데이터를 활용한 데이터 분석을 통해 기업 또는 조직의 의사 결정을 지원할 수 있는 비즈니스 인텔리전스에 관심이 있습니다. 이를 위해 다양한 데이터 분석 기법과 machine learning, deep learning을 공부하고자 노력하고 있습니다.
  - Keyword: Data analytics, Text mining, Machine learning, Deep learning, Business intelligence
- 

## 학력

- 2015.3 ~ 2020.2, 건국대학교 (졸업예정)
    - 공학사(산업공학)
- 

## 경력

- Business Intelligence & Data analytics Lab. @ Konkuk University (2018.04 ~ 2019.02)
    - 연구 인턴 ([Link](#))
  - Production Management Lab. @ Konkuk University (2019.03 ~ 현재)
    - 연구 인턴
- 

## 발표

- **김명섭**, 박영재, 이승주, 이권능, 최재은. (2019). 소셜 미디어 시계열 예측을 이용한 고객 니즈의 부상성 탐지: 워드 임베딩, 네트워크 분석, LSTM기반의 부상 키워드 탐지 방법. 제 15회 한국 대학생 산업공학 프로젝트 경진대회(대한산업공학회 추계학술대회), 서울대학교, 서울, 11월 8일. [은상 \(Link\)](#)
- **김명섭**, 민희식, 최재은, 이종민, 이승현. (2019). 온라인 데이터 마이닝을 활용한 서울지역 호텔의 외국인 고객 이용 만족도 요인 파악 및 경영 전략 개발. KMAC 경영혁신 연구논문 및 사례 연구 대학(원)생 공모전, 여의도 CCMM빌딩, 서울, 11월 2일. [장려상 \(Link\)](#)

- 박영재, **김명섭**, 엄기수. (2019). 소셜미디어 데이터 마이닝을 활용한 스마트 스피커 제품에 대한 기회 분석. 한국 산업경영시스템학회 대학생 프로젝트 경진대회, 공주대학교, 천안, 5월 3일. **우수상**. ([Link](#))
- 최재웅, **김명섭**, 윤장혁. (2018). 오픈데이터 분석적합성 평가 및 소셜 미디어를 활용한 비즈니스 인텔리전스 연구에 관한 문헌 분석. 대한산업공학회 추계학술대회, 한양대학교, 서울, 11월 9일. ([Link](#))

## 연구

- **소셜 미디어 시계열 예측을 이용한 고객 니즈의 부상성 탐지: 워드 임베딩, 네트워크 분석, LSTM기반의 부상 키워드 탐지 방법(2019.07 ~ 2019.10)**
  - 개인 연구 ([Link](#))
  - 제 15회 한국 대학생 산업공학 프로젝트 경진대회 본선 발표 예정(2019/11/8)
  - 소셜 웹 사이트 Reddit에서 subreddit이 Home automation(가정 자동화)에 속하는 게시글과 댓글 328,010건을 수집, 수집된 게시글과 댓글을 1개월 별로 분류한 후 띄어쓰기 단위로 분할, 이후 Word2vec을 사용하여 수집된 문서를 저차원으로 임베딩, 임베딩 차원에서 키워드 간의 거리를 계산하여 거리행렬을 생성한 후 가중치 식을 적용하여 키워드 간의 거리가 가까울수록 큰 값을 가지는 가중치 행렬 생성, 이후 문서로부터 동시 출현 행렬을 생성하여 가중치 행렬과 요소별 곱하여 단어 네트워크를 생성, 단어 네트워크의 근접 중심성을 키워드의 중요도로 간주, 2015년 12월부터 2019년 3월까지 총 40개월의 키워드 중요도를 input으로 LSTM을 사용하여 향후 12개월의 월별 키워드 중요도를 예측, 수집하여 계산된 40개월의 키워드 중요도를 현재 중요도, 예측된 12개월의 키워드 중요도를 미래 중요도로 간주, 미래 중요도를 현재 중요도로 나누어 부상성 계산, 현재 중요도와 부상성을 이용하여 포지셔닝 수행, 향후 중요도가 더욱 높아질 고객 니즈(키워드)를 탐지
- **온라인 데이터 마이닝을 활용한 서울지역 호텔의 외국인 고객 이용 만족도 요인 파악 및 경영 전략 개발 (2019.03 ~ 2019.10)**
  - 학사학위 청구논문 주제 ([Link](#))
  - KMAC 경영혁신 연구논문 및 사례연구 대학(원)생 공모전 본선 발표 예정(2019/11/2)
  - 국내 저널 투고 예정(저널 미정)
  - 트립닷컴 사이트에서 서울지역 호텔 데이터 443건 및 외국인 고객의 리뷰 데이터 110,517건을 수집, 외국인 만족도에 영향을 미칠 수 있는 약 148개 선택속성을 선정한 뒤 Vector Space Model을 사용하여 텍스트 데이터를 수치 데이터로 변환, 고객의

기대를 의미하는 가격과 만족을 의미하는 평점을 곱한 수치를 호텔의 효용으로 간주하여 종속변수, 148개 선택속성을 그룹화 하여 32개 변수를 선정 및 독립변수로 사용하여 multivariate regression을 수행, 효용에 영향을 미치는 선택속성을 파악, 이후 파악된 변수를 사용하여 계산한 서비스 특성 수준(Service attribute level)과 기대를 고려한 만족 수준을 두 축으로 사용하여 443개 호텔을 포지셔닝, 총 4개 영역의 호텔 중 선도 그룹에 해당하는 호텔을 다른 영역의 호텔들이 벤치마킹 할 수 있도록 키워드의 평점을 분석하여 경영 전략을 도출

• **소셜미디어 데이터 마이닝을 활용한 스마트 스피커 제품에 대한 기회 분석(2019.03 ~ 2019.05)**

- 연구 보조

- Google big query를 사용하여 reddit 사이트에서 경쟁 상태에 있는 스마트 스피커 제품 리뷰 데이터 약 20만건을 수집, LDA-topic modeling, sentiment analysis를 사용하여 제품 주제에 대한 중요도와 만족도를 계산, opportunity algorithm을 이용하여 중요도와 만족도를 바탕으로 한 제품 토픽의 기회를 도출, 제품 토픽의 기회를 비교 분석

- 기여 내용: 선행 연구에 대한 문헌 분석, topic labeling, 결과 해석

• **특허의 토픽 모델링과 link prediction을 이용한 기술주제 융합전망 탐지(2019.03 ~ 2019.05)**

- 연구 보조

- IPC 분류코드가 H04(Electric communication technique)로 시작하는 66,029개 특허를 대상으로 특허 초록에 LDA-topic modeling을 수행, document-topic matrix에서 특정 threshold를 기준으로 문서에 속하는 topic선정, 문서를 기준으로 한 topic의 동시 출현 관계를 파악, 이후 link prediction을 수행하여 현재 한 특허 내에서 빈번히 동시 출현하지 않지만 향후 빈번히 동시 출현할 topic의 쌍과 확률을 탐지

- 기여 내용: 방법론 설계, document에 속하는 topic 결정, topic co-occurrence matrix 작성

• **그래프 기반 엔트로피를 이용한 소셜 미디어와 특허의 동향 분석(2018.9 ~ 2019.2, 미완)**

- 개인 연구

- 스마트폰 특허의 초록 및 소셜미디어 댓글에 대하여 Graph-based entropy를 사용하여 기술을 대변하는 특허의 토픽과 소비자를 대변하는 소셜 미디어 토픽의 시계열적 분화, 합산, 변화를 파악하고자 함

- 진행 사항: 데이터 수집, keyword 추출, LDA-topic modeling, co-occurrence 기반 graph 작성

• **근미래이슈 탐지를 위한 오픈데이터 유형탐색 및 분석모델 연구(2018.05 ~ 2018.10)**

- 한국과학기술정보연구원(KISTI) 위탁과제
- 학생 연구원으로 참여
- 급변하는 비즈니스 환경의 외부적 요인을 연구에 반영하고자 다양한 오픈데이터를 탐색하고 분류하여 계층구조를 정리, 오픈데이터의 분석적합성을 평가하고 실제 운영되고 있는 사례를 탐색, 오픈데이터를 활용한 주요 논문을 검색 및 선별하여 구체적인 분석방법과 사용한 데이터를 분석
- 기여 내용: 선별된 문헌에 대한 문헌연구. 사용한 오픈데이터의 유형 및 오픈데이터 출처, 사용된 방법론 및 세부사항과 연구의 기여점을 요약

---



---

## Skill

- Programming – Python
- Statistics – SPSS, Minitab
- Database – MySQL
- Simulation – Arena

---



---

## Data science study

- Business intelligence & Data analytics Lab. Machine learning study (2018.07 ~ 2018.08)
  - NLP(Word2vec) 주제로 발표 [\(Link\)](#)

---



---

## Edwith online lecture

- 머신 러닝을 위한 python 워밍업, 최성철 – 수강
- 신경망과 딥러닝, Andrew Ng – 수강 중
- 자료구조 및 알고리즘 개론 1, 문일철 – 수강 중

---



---

## 학부 프로젝트

- 2018. spring. Management of technology (9팀 중 1위)
  - I-sence Co. R&D planning with e-skin technology [\(Link\)](#)
  - **김명섭**[리더] 외 3명
  - 기업의 보유기술을 기반으로 short-term business와 long-term business에 적용할 수 있는 두 개의 제품을 개발하기 위한 기술의 발굴과 습득 및 제품 개발을 위한 전략 구상, Technology opportunity discovery, 5 forces model, Social media mining, STP analysis, Diffusion model, TRM, Leverage analysis, Disruptive innovation, Chasm management, Business model canvas등의 방법 이용

- 2018. spring. Data analytics (9팀 중 3위)
  - 합격한 자기소개서의 내용적 특성을 파악하기 위한 자기소개서 text mining ([Link](#))
  - **김명섭**[리더] 외 4명
  - 기업 합격에 영향을 미칠 수 있는 자기소개서의 내용적 특성을 파악하기 위해 합격한 자기소개서와 합격 여부를 알 수 없는 자기소개서를 수집하여 keyword의 tf, tf-idf, sentiment를 비교 분석하고 document의 personality insight를 비교 분석
- 2018. fall. Enterprise DB modeling (7팀 중 1위)
  - 통합적 공연관리 플랫폼을 위한 데이터베이스 구축 ([Link](#))
  - **김명섭**[리더] 외 5명
  - 문화예술공연 분야의 VOC를 수집하여 주체가 되는 공연자, 공연장, 관객의 관점에서 각 주체들의 불만사항을 조사 후 현재 서비스되고 있는 플랫폼이 VOC를 만족시키지 못하는 한계를 극복하기 위해 필요한 기능을 정의하여 관계형 데이터베이스를 구축
- 2018. fall. 시뮬레이션 (7팀 중 1위)
  - 응급실 프로세스 모델링 ([Link](#))
  - **김명섭**[리더] 외 2명
  - 주어진 조건에 따라 응급실의 프로세스를 모델링하고 고객에게 발생할 수 있는 불편사항을 최소화할 수 있는 척도를 가치를 부과하는 시간, 가치를 부과하지 않는 시간, 전체 프로세스의 시간 세 분류로 설정, 통제가 가능한 요소인 침상 수, 의사 수, 직원 수를 조정하여 고객의 만족도와 병원 응급실 경영에 있어 모두 최적이 되는 대안을 발견.

---

## 학부 topic presentation

- 2018. fall 기술혁신론
  - 선정된 주제에 대하여 학생이 직접 자료를 준비하여 1시간 동안 강의
  - Almost everything of amoeba management (topic: 아메바 경영) ([Link](#))
    - 아메바 경영의 방식, 독립채산제도, 철학, 리더십 등을 강의
  - 미래기업의 조건 – Strategies for Survive (topic: 파괴적 혁신) ([Link](#))
    - 파괴적 혁신에 대한 소개, 변화의 신호 등을 강의

---

## 자격증

- 정보처리기사 (자격번호: 19202021302P)
- ERP 생산 정보관리사 1급
- ERP 물류 정보관리사 1급
- SMAT 서비스경영 1급