

# 텍스트 마이닝

마틴 루터 킹 연설문과 실시간 뉴스 분석

제출일	2021, 04, 08
작성자	정승호

## 목 차

서론	3
본론	
1. 마틴 루터 킹 연설문 토픽 분석	3
2. 연설문 분석결과 연관분석	4
3. 실시간 뉴스데이터 토픽분석	5
결론	6

■ 서론

이 레포트는 1963년 8월 28일 마틴 루터 킹 박사가 연설한 “I have a dream”의 국문 번역문을 바탕으로 텍스트 마이닝을 실시한 결과와, 실시간 뉴스 데이터를 텍스트 마이닝한 결과를 다룬다.

■ 본문

- (1) 데이터를 이용하여 토픽 분석을 실시해 단어구름으로 시각화 하고 단어 출현 빈도수를 기반으로 어떤 단어들이 주요 단어인지 분석

단어구름 시각화

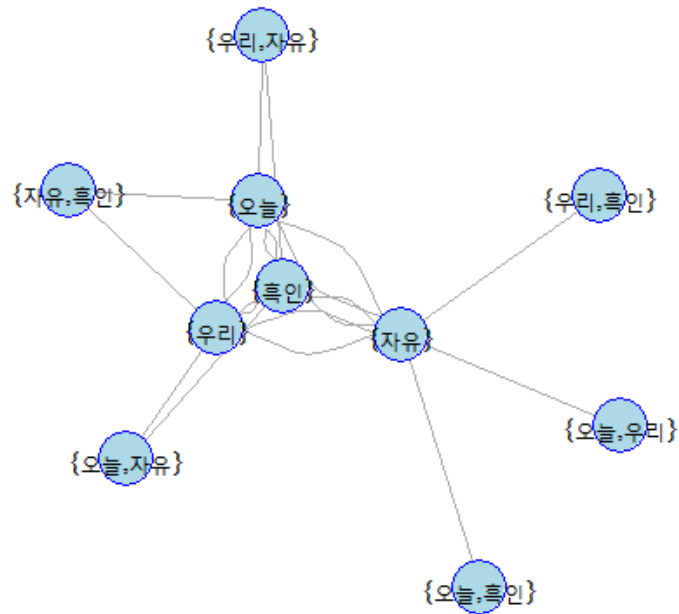


분석결과

단어구름 생성결과와 데이터 정제결과 확인을 통해 해당 연설문에서 가장 중요했던 단어는 '우리(맥락상 흑인과 동의어)'(33회), '자유'(27회), '흑인'(19회) 등임을 확인할 수 있으며, 이를 통해 당시 차별당했던 흑인들의 자유와 해방을 원했던 글쓴이의 의도를 파악할 수 있다.

(2) 데이터를 이용하여 연관분석을 실시하고 시각화하여 분석 결과 확인.

## 연관분석 시각화



## 분석결과

설명3: 연관어 분석을 통해 '우리', '흑인', '오늘', '자유'등의 키워드가 가장 핵심임을 확인할 수 있으며, 이들은 {오늘, 자유}, {우리, 자유}, {오늘, 흑인}, {우리, 흑인}, {오늘, 우리}, {자유, 흑인}의 형태로 연관되어 있다. 앞서 살펴본 단어구름과 연관어 분석을 통해 산출된 결과를 통해 연설문에서 강조했던 단어들인 '우리', '흑인', '오늘', '자유'를 확인할 수 있다.



## ■ 결론

이 레포트를 통해 1963년 8월 28일 마틴 루터 킹 박사가 "I have a dream"이라는 제목의 연설문의 국문 번역문의 단어 빈도분석과, 연관어 분석, 실시간 뉴스 분석을 R code를 이용해 텍스트 마이닝을 제대로 학습할 수 있었다.