따릉이는 서울시가 주관하는 무인 공유 자전거 대여 사업이다. 2015년 10월 2,000대 공급을 시작으로 2019년 30,000대 공급을 목표로 하고 있다. 2017년 12월에 약 20,598대의 공공자전거가 운영되면서 인구 1만 명당 20대의 공공자전거 운용이라는 초기의 목표를 조기 달성했다. 그 결과 2018년 5월 기준 총 누계 이용 건수 2,393,531건, 총 회원 수는 753,109 명이 사용하며 시민의 발이 되어주는 중요한 공공 서비스로 자리 매김하게 되었다.(서울시 연구원 – ‘공공자전거 종합현황’ )

따릉이 공공 자전거 서비스를 사용하기 위해서는 어플리케이션 혹은 웹페이지 사용한 예약이 필수적이다. 하지만 서울시가 조사한 만족도 대비 2019년 7월 기준 50만 다운로드 횟수를 기록한 따릉이 의 ‘Google Playstore’ 어플리케이션의 점수는 1.59 점(총 리뷰수: 1,847개)을 기록하고 있다. 또한 서울시 연구원에서 조사한 바에 따르면 2018년 5월기준 홈페이지(시민의 소리 게시판)에 관련된 누계 불만 처리 수는 2,737건 이지만, 어플리케이션과 웹페이지에 관련된 불만 사항 갯수는 76개 밖에 되지 않는다**(실제 2018년도 낮은 별점을 단 데이터 갯수 추가).** 이를 보아 따릉이를 이용하는데 있어 불만 사항 많은 의견들이 충분한 피드백을 받지 못하는 것을 알 수 있었다.

그러므로 본 연구를 진행 하기에 앞서 ‘Google Playstore’ 상에서의 따릉이 어플리케이션 리뷰와 홈페이지의 시민의견수렴 게시판의 불만사항 및 문제점을 분석을 진행했다. 설문 조사의 경우 샘플 수가 상대적으로 적으며, 사전에 문항을 선정하기 때문에 이용자들의 다양한 경험을 반영하기 어렵다는 한계가 있다.(**어떤 문제 점이 있는지 알기 위해서 데이터 분석을 진행 했다 라고 수정)** (Coughlan et al., 2009). 구글 플레이 스토어에 기재된 리뷰와 시민의견 수렴 게시판의 경우 설문 조사에 방대한 데이터를 수집 할 수 있었고, 실제 어플리케이션을 사용한 사용자의 경험을 반영하고 있다. 따라서 데이터 마이닝 기법을 활용하여 설문 조사를 사용자의 불만 사항과 문제점을 분석하였다.

먼저 텍스트 분석을 하기에 앞서 2015년 9월 부터 2019년 7월 까지 ‘Google Playstore’에 등 록된 ‘서울 자전거 따릉이’ 리뷰 그리고 2016.11.13 부터 2019.08.03 까지 따릉이 홈페이지 ‘시민의견수렴’ 게시판에 게시된 모든 게시글들을 파이썬 라이브러리 ‘Beautifulsoup 4’ 와 ‘Selenium’ 을 사용하여 크롤링(Crawling) 하였다. 수집 항목은 다음과 같다. ‘Google Playstore’의 경우 평점, 이용자명, url 항 목을 수집하였고, 시민의견 수렴 게시판은 작성자 ID, 게시날짜, 제목, 내용으로 수집을 진행했다.

인터넷 상에서 수집된 정보는 띄어쓰기 및 맞춤법 오류, 고유명사의 다양한 사용 등 여러 예외적 사용이 많기 때문에 텍스트 분석을 진행하기에 앞서 데이터 전처리가 필요하다. 데이터 전처리를 위해 파이썬 패키지 중 KoNPLPy Twitter와 KoNLPy의 수정된 버전인 customized KoNLPy (<https://github.com/lovit/customized_konlpy>) 를 사용하였다. customized KoNLPy은 기존 koNLPy에서 할 수없었던 단어(사전) 추가 기능과 검색어 제외, 품사 추가 기능을 제공한다.

기본적으로 단어를 기반한 문제점을 찾기 위해 명사, 형용사를 추출하여 빈도수를 계산하였다.

분석된 결과는 다음과 같다.

이러한 사용 형태를 보아 실사용자의 인식 혹은 평가를 살펴보기 위해 ‘서울자전거 따릉이’ 어플리케이션의 사용자 리뷰를 살펴 보았다. 또한 웹페이지를 통한 사용자의 불만 사항 혹은 사용자의 의견을 파악 하기 위해 ‘서울자전거 따릉이’ 사이트에 속해 있는 시민의견 수렴 게시판의 글들을 텍스트

텍스트 마이닝 분석은 다음과 같다.