

장애인 이동권 보장을 위한 LSTM 기반 장애인 콜택시 대기시간 예측 연구

이정민*·박영빈**

현재 장애인의 이동권 문제는 여전히 해결되지 못한 실상이다. 그중에서도 장애인 콜택시 게시판 이용자들의 후기글을 워드클라우드 및 의미 연결망 분석 시각화한 결과, 장애인 콜택시 대기시간 및 배차 시스템에 대한 불편사항이 있음을 파악하였다. 해가 지날수록 콜택시 대기시간은 감소하는 추세이지만 평균 30분을 기다려야 하며 최대 2시간 이상이다. 또한 대부분의 지역은 실제 콜택시 운영 대수가 법정 대수를 만족하지 못하는 실정이었다. 본 연구에서는 공공데이터를 수집 및 분석하여 장애인 콜택시 이용자들이 어떤 부분에서 서비스의 불편함을 느끼는지, 서비스 운영 상태가 어떠한지에 대해 파악하였다. 그리고 LSTM 기반 모델을 이용해 장애인 콜택시 대기시간을 예측할 수 있는 모델을 구현하였다. 에포크(epochs)를 4로 설정하였을 때, MSE가 0.9955로 가장 작은 값을 보여 최종 모델로 선정하였다. 이를 통해 실제 장애인 콜택시 대기시간을 감소시킬 수 있는 방안을 제안하였다. 더 나아가 우리 사회에서 장애인들의 삶의 질이 더욱 개선되고, 궁극적으로 이동권이 온전히 보장되길 기대한다.

주제어 : 장애인 이동권, 장애인 콜택시, 텍스트 마이닝, 대기시간 예측, LSTM

* 경찰대학 범죄학과 석사과정

** 국립강릉원주대학교 정보통계학과 학부

I. 서론

장애인의 인권과 삶의 질 개선을 위한 노력은 계속되고 있지만 여전히 부족한 상태이다. 그중에서 현재 국내에서 장애인과 관련된 가장 뜨거운 화두는 ‘이동권’임이 명백하다. 2018년 “장애인 사회 및 문화·여가 활동 실태와 정책과제”에 따르면, 집 밖 활동이 불편하다고 느낀 응답자는 46.6%이며, 장애인 교통수단 이용 시 36.8%가 어려움을 느낀다고 응답하였다. 전국장애인차별철폐연대(이하 전장연)는 현재도 장애인 이동권 보장을 위한 시위를 지속하고 있다.

장애인의 이동권 보장을 위해 지자체별로 교통약자 이동지원센터에서 장애인 콜택시를 운영하고 있다. 해가 지날수록 장애인 콜택시의 평균 대기시간은 감소하는 추세이고 운행 차량 수는 증가하고 있지만, 여전히 대기시간은 길고 배차 시스템에 대한 불편 사항이 존재한다. 서울시설공단 장애인 콜택시 운영처에 따르면, 2021년 장애인 콜택시의 평균 대기시간이 가장 긴 시간대는 오후 4시(평균 58분), 오후 5시(평균 32분), 오후 9시(평균 48분)로 나타났다. 장애인 콜택시 이용자들은 대기시간이 얼마나 오래 걸릴지 모른다는 점과 이로 인해 장애인 콜택시로 출퇴근이 어렵다는 점이 현 문제점이다. 또한 저녁 시간대에는 운행 차량이 크게 줄어들어 대기시간이 증가한다. 일부 지자체(서울)의 교통지원 어플리케이션(이하 앱)에서는 평균 대기시간을 제시하지만, 그렇지 않은 지자체가 더 많았다. 서울시각장애인 생활·이동지원센터 운영의 자료(22.03.21~22.03.25)에 의하면 하루 평균 접수 건수는 1,348건, 탑승 건수는 1,051건(탑승률: 78%)으로 집계되었다. 30분 이내 배차 완료는 830건(79%), 30분 이상 1시간 이내 배차 완료는 168건(16%), 1시간 이상 2시간 이내 배차 완료는 51건(4.9%)이었다. 접수 2시간이 지나면 자동으로 접수가 취소된다. 이러한 이유로 장애인 콜택시 이용객들은 대기시간이 오래 걸려 접수를 취소하고, 다시 시도한 경우까지 고려하면 배차를 받는 데에 매우 많은 시간이 소요된다. 게다가 배차가 완료된 뒤, 장애인이 실제로 차량에 탑승하기까지 평균 19분이 추가로 소요된다. 따라서 위의 상황을 고려해보면, 장애인이 장애인 콜택시를 배차 신청한 후에 배차 신청이 완료되며, 실제로 탑승하기까지는 많은 시간이 소요될 것으로 예상된다. 이처럼 장애인은 비장애인에 비해 이동수단을 타기 위해 적게는 4배, 많게는 6배까지 대기하고 있어야 하는 것이 현실이다.

이러한 현황은 장애인 이동권의 문제점을 여실히 보여준다. 따라서 본 연구에서는 장애인 이동권 보장을 위해 장애인 콜택시에 초점을 맞춰 콜택시 이용자들이 서비스를 이용하는 데에 겪는 문제점을 빅데이터 분석으로 이해하고, 이들의 불편을 줄일 수 있는 방안을 제시하고자 한다. 또한 이 연구를 통해 자유로운 취업 및 여가활동을 도와 자립할 수 있는 환경을 추구함으로써 장애인들의 삶의 질을 개선하고, 궁극적으로 이동권이 온전히 보장되는 데에 도움이 되고자 한다.

II. 선행연구 고찰

1. 특별 교통수단 불균형 개선 연구

이병화·양희택(2017)은 장애인 콜택시 광역이동지원센터를 운영하고 있는 서울시 운영 자료와 경기도의 현 실태 자료 분석을 통해 센터 개선 사항을 점검하고, 후발 광역 이동 지원 센터 설립을 위한 대안을 마련하는 연구를 진행하였다. 서울특별시 장애인 콜택시 운행 빅데이터와 경기도 장애인 콜택시 운행 및 이용자 데이터를 활용하여 출발지역과 목적지역 등의 운행지역, 이용목적, 장애유형 등 이용자 특성, 탑승 소요시간과 배차지연 시간과 이유 등을 빈도 분석하였다. 관내 이동의 경우 저상버스 등을 개발하여 출발지와 목적지를 다양화하고, 장애인 콜택시의 이용목적 확대시키고 탄력적으로 운영할 필요성, 휠체어 여부에 따라 다양한 장애인 콜택시 유형의 다변화, 지연시간의 단축 방안 필요성을 제언하였다. 석종수(2014)는 인천광역시 장애인 콜택시 콜센터의 전산 자료를 빈도 분석하여 이용자의 장애 유형을 파악하고, 행정동별 호출 건수를 비교하였다. 또한 이용목적별 운행 비율과 시간대별 이용 건수, 장애인 콜택시 대기시간의 분포를 시각화하여 대기시간을 줄이기 위한 방안을 강구해야 한다고 제언하였다. 성광환·신유열·윤윤진(2022)은 서울시설공단의 장애인 콜택시 데이터 셋을 활용하여 서울시 지역별 장애인 콜택시의 접근 형평성이 어떠한지 공간회귀분석을 진행하여 시간적, 공간적 접근성의 차이가 존재함을 증명하였다. 장애인 남성 비율과 종사자 밀도 비율이 높을수록 콜택시의 접근성이 좋지 않았고, 단위 면적 당 버스 정류장이 많을수록, 기초수급대상자 비율이 높을수록 접근성이 좋은 편이었다. 또한 불균형을 제거함으로써 지역 전체의 접근성을 향상시키는 운영을 제안하였다. 정진형·박호성·여윤구·김학빈·강준혁(2018)은 2017년 기준으로 장애인 콜택시 보급률이 가장 적은 전라남도 지역을 대상으로 고객의 대기시간과 비용을 고려한 최적의 할당 대수를 결정하는 방향을 제시하였다. 우선 AHP 분석¹⁾으로 장애인 콜택시 할당 우선도시를 선정하였고, Queueing Theory(대기이론)를 활용하여 도시 별 대기시간 산출 및 최적의 차량 대수를 결정하였다. 또한 수도권 지역에 비해 타 지역의 구체적인 DB 부족 문제점을 지적하며 체계화된 DB 구축의 필요성을 제기하였다. 신용은·최혜미·송기욱·이희대(2014)는 경상남도 콜센터 자료를 활용하여 특별교통수단 이용자의 통행 특성을 파악하였다. 수집된 자료에 기록되는 항목은 출발지, 도착지, 장애 종류, 휠체어 이용 여부, 배차 결과 및 실패 사유 등을 포함한 총 10개 항목이고, 이를 가공하여 요인별로 분석하였다. 휠체어 이용자는 42%이며 주 이용자는 신체 및 뇌병변 장애인으로

1) Analytic Hierarchy Process(계층적 분석과정), 의사결정의 전 과정을 여러 단계로 나눈 후 일을 단계별로 분석 해결함으로써 합리적인 의사결정에 이를 수 있도록 지원해 주는 방법

나타났지만, 노인과 임신부의 비율은 낮았다. 주중 및 주말 이용 패턴은 일반인과 유사하지만 시간대별 침두시간은 일반인의 비침두시간에 집중됨을 확인하였다. Feeley, C., Spasovic, L. N., Dimitrijevic, B., & Lee, J.(2016)는 장애인 대중교통 승객을 위한 차량 도착 알림 시스템을 제작하였다. 해당 시스템은 Google Traffic API를 이용해 차량을 추적하며 3G 및 4G-LTE를 활용하도록 설계되었다. 또한 승객 알림 외에도 Google Maps API를 통해 개발된 실시간 경로 안내 정보를 제공한다. 이를 통해 대기시간은 기존 20~40분에서 15~20분으로 줄어들었고, 거의 모든 경우에 차량이 예정 도착 시간보다 1분 일찍 도착함을 확인하였다.

이러한 연구들을 통해 지역별로 장애인 콜택시 서비스가 접근성 측면에서 불균형하게 이루어지고 있다는 현황을 파악할 수 있었으나, 어떠한 방식으로 지연 시간을 단축시킬 수 있을지 후속 연구가 필요할 것이다.

2. 특별 교통수단 배차 최적화 연구

홍두선·가은한·하성현·이청원(2018)은 서울시 장애인 콜택시 대기시간 예측 LSTM 모델의 적정 하이퍼 파라미터 선정에 관한 연구 실시하였다. 하지만 더 많은 하이퍼 파라미터를 고려할 필요가 있으며, 야간 시간대 특성상 적은 탑승 건수로 야간 시간대의 정확도가 낮은 한계가 있다고 밝혔다. 박형수·김훈기·김영성(2017)은 교통약자를 위한 시간 복잡도를 반영한 자동 배차 알고리즘에 대한 연구를 실시하였다. 자동 배차 기능의 문제점을 발견하여 최선의 자동 배차 알고리즘 구현하였으나, 도/전국 단위로 시스템을 구축하지 않아 효율적인 시간 복잡도를 갖는 알고리즘을 구현하는 데에 한계점이 있다. 손중훈·김도경(2021)은 생존분석을 활용한 장애인 콜택시 이용자 대기시간 영향요인에 대한 연구를 실시하였다. 장애인 콜택시 이용자가 호출 후 승차까지 소요되는 시간에 영향을 미치는 영향의 정도를 규명하기 위해 Cox 비례위험모형, 3가지 공변량(운전원 출퇴근 시간대, 장애인 콜택시 회원 규모, 차량 이용 가능도)을 활용하여 제시하였다. 평일과 주말에 따라 대기시간에 영향을 미치는 효과가 다른 것과 요일별 특성이 같더라도 호출 시간대 및 지역 등의 시공간적 특성에 따라 이용자가 경험하는 대기시간의 증감률이 다른 것을 확인하였으나 단지 3가지 공변량만 고려한 것에 한계점이 있다. 박진혁·장영재(2015)는 서울시 장애인 콜택시 고객 대기시간 감소를 위한 자동 배차 알고리즘 및 최적 차량 공급을 제안하였다. 콜 접수 내역(접수 시각, 희망 탑승 시각, 고객 유형, 등), 실시간 차량 데이터(위치, 상태, 등), 운전자 근무 배치표(운전자 출퇴근 시간, 등) 데이터 분석을 통해 고객 대기시간 세분화 및 배차 알고리즘 요건 파악하였다. 그리고 배차 최적화 모델을 구현하여 실질적인 공급 차량 판단 기준 확립하여 공급 최적화 모델 구현하였다. 김건욱·윤대식·김종진(2020)은 빅데이터를 활용하여 대

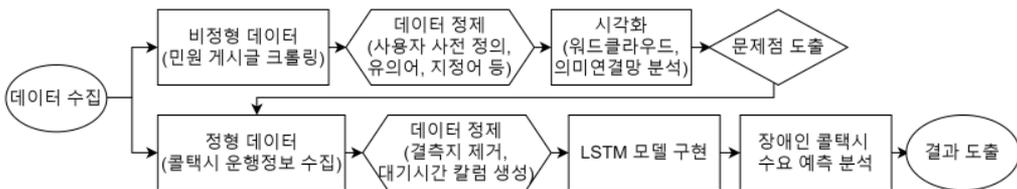
구광역시의 교통약자의 특별 교통수단 통행 수요를 분석하였다. 특별 교통수단 이용자들의 실제 배차 이력 자료(특별 교통수단 이용자 수, 콜센터 직원 수, 운행 차량 대수, 등)와 기상자료(온도, 강우량, 미세먼지, 등)를 활용하여 이용자들의 통행 수요를 시점별 통행 특성(요일별, 시간대별, 통행 시간 및 대기시간, 등)과 공간적 통행 특성(승하차 지점, 승하차 지점별 대기시간, 등) 두 가지로 나누어 분석하였다. 또한 기상 상황별(미세먼지, 적설 여부, 강우 여부, 온도) 특별 교통수단 이용자 수를 분석하였으나 시간적 범위의 제약으로 인하여 정확한 비교 분석이 어렵다는 한계점이 존재한다. 이도엽·정수훈·천영재·김동일·유승일(2018)은 택시 수요 예측 모델인 TGNet을 개발하였다. TGNet은 카카오브레인과 카카오모빌리티가 개발한 모델로 단순한 시계열 예측이 아닌 시공간 상의 수요 예측 모델이다. 이재길(2020)의 연구에 따르면 TGNet 모델은 여러 시간대의 과거 택시 수요를 여러 채널의 이미지로 간주하여 CNN만을 사용하였고, 시간 맥락을 고려하기 위해 시간 정보를 연속된 실수 값으로 표현하는 Embedding layer를 추가하였다. 또한 택시 하차와 같은 보조 정보를 고려할 수 있도록 추가 임베딩을 고려하였다. 과거 4시간의 8개 채널의 데이터, 시간 임베딩 결과, 추가 임베딩 결과를 결합하고 Fully-Connected 층을 통과시켜 다음 시간대의 택시 수요를 예측하게 된다. 교통량, 날씨와 같은 외부 데이터를 사용하지 않고도 높은 성능을 보여주어 기존 연구와 차이점을 보인다. Garnier, c., Trepanier, M., Morency, C.(2020)는 대중교통 차량이 승객들을 탑승시키고 하차시키기 위해 정차하는 시간인 체류시간을 종속변수로 선정하고, 차종, 정류장 위치, 날짜 및 시간, 휠체어 이용 여부와 같은 승객 특성을 독립변수로 하는 선형회귀모형을 구현하였다. 이로 인해 경로가 최적화되었고, 운영비용을 줄이는 데에 도움이 되는 결과를 보였다.

앞선 선행연구를 통해 장애인 콜택시 데이터의 한계로 이동 거리와 이동 시간과 같은 주요한 변수가 모델에 포함되지 못하여, 대기시간을 정확히 예측하는 데 어려움이 있다는 것을 확인할 수 있었다. 또한 도착시간(운행 완료시간)을 제공하지 않아 가용 가능한 차량에 대한 정보가 없다는 것이 모델의 예측력에 영향을 미쳤을 수 있다고 판단되었다. 더불어 시군구와 같은 공간적 정보가 반영되지 않아 지역별 수요를 정확하게 파악할 수 없어 추후 공간적 정보 등과 같은 데이터를 확보할 수 있는 방안이 필요할 것으로 사료된다.

Ⅲ. 연구 범위 및 절차

1. 연구 방법

본 연구는 아래 순서도와 같이 진행하였다. 장애인 콜택시 이용자들의 서비스 이용 후기를 알고자 서울시설공단 - 장애인 콜택시 페이지의 민원 게시글을 웹 크롤링(Web crawling) 기법으로 수집하였다. 텍스트 데이터는 글을 쓰는 사람의 표현 방법과 글을 작성하는 당시의 기분 상태에 따라 단어 선택 및 어휘 표현이 상이하므로 이를 통일하기 위한 정제 작업을 진행하였다. 그 후, 정제된 데이터를 이용해 워드클라우드와 의미연결망 분석 기법으로 장애인 콜택시를 이용하는 장애인들의 불편사항을 파악하였다. 도출된 결과에 대해서 면밀히 탐색하고자 “서울 열린데이터광장”에서 서울시 장애인 콜택시 이용 현황을 알 수 있는 장애인 콜택시 운행 정보에 대한 Open API 데이터를 데이터프레임 형식으로 수집하였다. 2021년 데이터를 대상으로 결측치를 제거하였고 모델링에 필요한 파생변수를 생성하는 등의 작업을 수행하였다. 마지막으로 앞서 생성한 파생변수를 대상으로 LSTM 알고리즘을 이용해 장애인 콜택시의 수요를 예측할 수 있는 방안을 제시하고자 하였다.

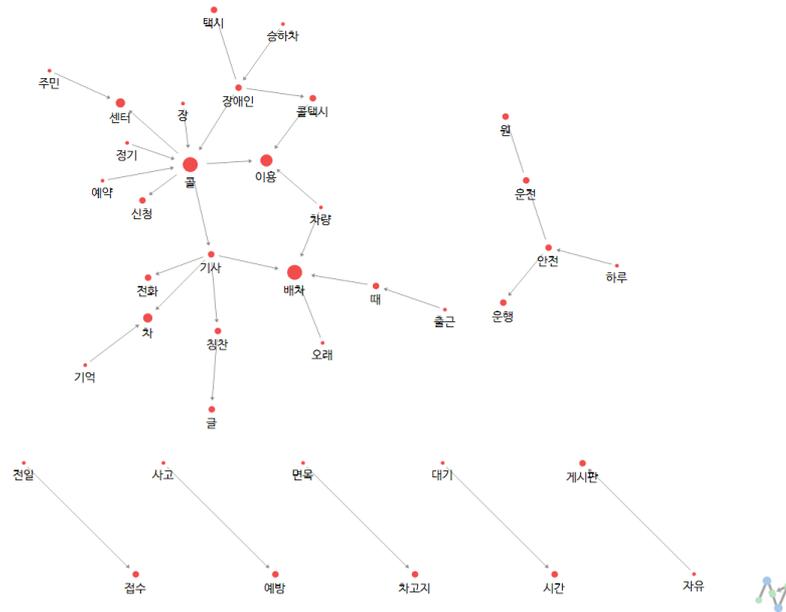


〈그림 1〉 연구 순서도

2. 활용 데이터 소개

1) 비정형 데이터 수집 및 시각화

우선 장애인 콜택시를 이용한 사람들의 인식을 알기 위해 [서울시설공단]-[장애인 콜택시]의 [시민의 소리] 게시글을 수집하여 장애인 콜택시를 이용하는 장애인들의 불편사항을 파악하였다. 2019.01.07 ~ 22.06.30 기간 동안 작성된 게시글의 작성일, 제목, 내용을 Jupyter Lab 환경에서 Python의 urllib, selenium 패키지를 이용해 크롤링하여 데이터를 수집하였다. 다양한 지역의 의견을 수렴하고 싶었으나, 이러한 이용 후기 게시글을 공개하는 창구는 서울시설공단이 유일

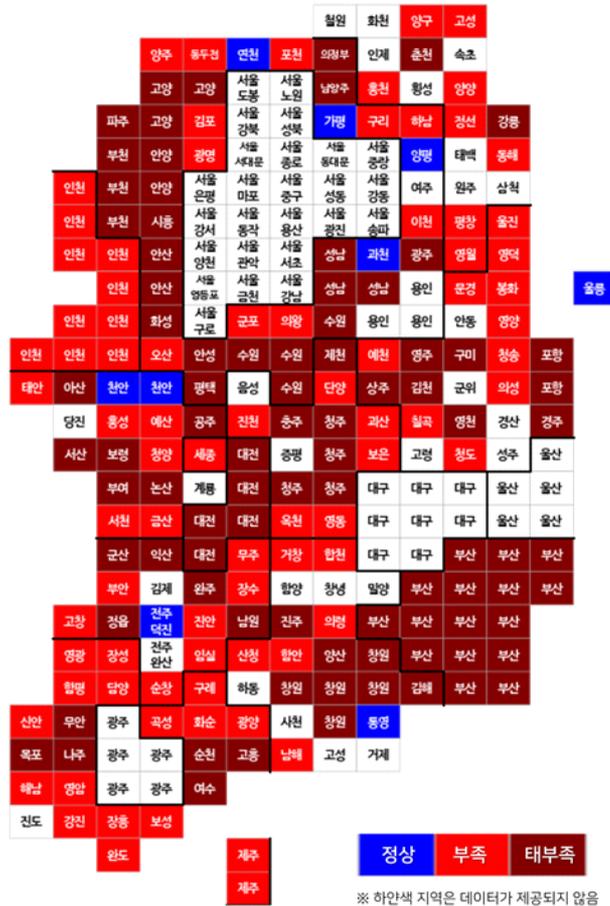


〈그림 3〉 의미연결망 시각화

또한 교통약자 이동편의 증진법²⁾에 따르면 중증장애인 150명당 1대콜로 장애인 콜택시가 도입되어야 한다. 각 지역별로 실제 장애인 콜택시 운영 대수가 법정 대수를 만족하고 있는지 확인하기 위해 공공데이터포털의 “전국 교통약자 이동지원센터 정보 표준 데이터”와 국가통계포털의 “시군구별, 장애정도별, 성별 등록 장애인수” 데이터를 활용하였다. 법정 대수에서 운영 대수를 뺀 차이를 기준으로 차이가 양수이면 ‘정상’, 중앙값(median)보다 차이가 큰 지역은 ‘태부족’, 그 외에는 ‘부족’으로 표시하였을 때 지역별 심각성을 명확하게 파악할 수 있다. 앞서 정의한 정상, 부족, 태부족을 지도 시각화하였으며, 색이 없는 흰 부분은 서울시와 일부 광역시 및 지역으로 실제 콜택시 운영 대수와 같은 정보가 존재하지 않아 현황을 확인하는 것이 불가능하였다. 해

2) 교통약자의 이동편의 증진법 시행규칙 [시행 2019. 7. 5.] [국토교통부령 제637호, 2019. 7. 5., 일부개정] 제5조 제1항을 다음과 같이 하고, 같은 조 제2항 중 “「장애인복지법」 제32조에 따라 등록된 제1급 및 제2급 장애인”을 “「장애인복지법 시행규칙」 제28조 제1항에 따른 보행상의 장애인으로서 같은 규칙 별표 1에 따른 장애의 정도가 심한 장애인”으로 한다.
 ① 법 제16조 제1항에서 “국토교통부령으로 정하는 대수”란 「장애인복지법 시행규칙」 제28조 제1항에 따른 보행상의 장애인으로서 같은 규칙 별표 1에 따른 장애의 정도가 심한 장애인 150명당 1대를 말한다. 제6조 제1항 제1호 중 “「장애인복지법 시행규칙」 제2조 제1항에 따른 1급 또는 2급 장애인”을 “「장애인복지법 시행규칙」 제28조 제1항에 따른 보행상의 장애인으로서 같은 규칙 별표 1에 따른 장애의 정도가 심한 장애인”으로 한다.

당 정보가 존재한다면 실제 운영 대수가 법정 대수를 현저히 만족하지 못하는 지역을 우선하여 차량 부족의 문제를 해결할 수 있을 것으로 사료된다. <그림 4>를 통해 지역별로 실제 운영대수가 법정 대수를 만족하고 있는지 확인할 수 있다.



<그림 4> 전국 장애인 콜택시 운영 대수 지도 시각화

위의 결과들을 종합적으로 파악하여 장애인 콜택시 이용자들이 배차와 대기시간에 불편을 겪고 있음을 확인하였다. 이어서 정형 데이터 분석을 통해 콜택시 운영 현황과 이용자들의 대기시간을 파악하여 지역별로 어느 정도 대기시간이 소요되는지 파악하였다.

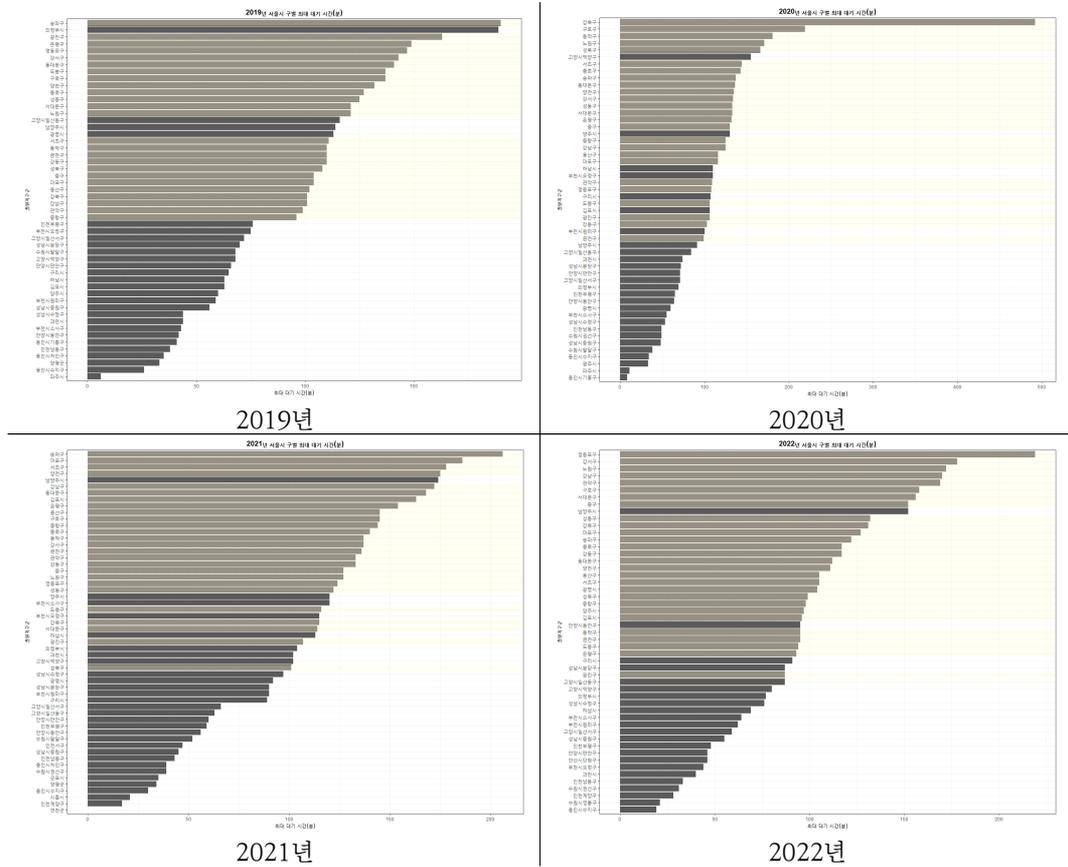
2) 정형 데이터 수집 및 EDA

2019.01.01 ~ 2022.08.20 기간의 서울시 장애인 콜택시 운영 현황을 파악하고자 서울 열린 데이터광장의 장애인콜시스템 Open API 데이터를 활용하였다. 데이터는 차량고유번호, 차량타입, 예정일시, 배차일시, 승차일시, 출발지, 목적지 정보로 구성되어 있다. Python에서 urllib, BeautifulSoup, datetime 등의 패키지를 활용하여 제공되는 XML 데이터를 파싱하고 데이터프레임으로 정리하였다. '예정일시'는 차량 승차 시점을 미리 설정한 시각, '배차일시'는 차량을 배차 받은 시각, '승차일시'는 실제로 차량에 탑승한 시각을 의미한다. 이후 데이터 정제는 R 프로그램을 활용하였다. 시간 계산의 용이성을 위해 오전, 오후로 나누어진 시간을 0 ~ 23시간으로 변경하였다. 예를 들어 '2000-01-01 오후 1:00:00'를 '2000-01-01 13:00:00'로 바꾸었다. 또한 예정일시, 배차일시, 승차일시는 문자형으로 수집되었기에 as.POSIXct(), strptime() 함수를 이용해 날짜 형으로 데이터 형 변환을 수행하였다. '대기시간' 칼럼은 (승차일시) - (배차일시), '분단위로 계산하였으며, 이 중에서 배차일시가 승차일시보다 이른 시각이 있어서 음수 값이 나오거나 비어 있는 값이 있으면 제외하였다.

이후에 연도별 출발지구군을 기준으로 최대 대기시간이 몇 분인지 정리하여 막대그래프 구현하였고, <표 1>과 <그림 5>에서 확인할 수 있다. 목적지가 서울권이면 인천과 경기도지역이 출발지에 포함되었다. 출발지구군이 서울 지역이면 형광색으로 표시하였다. 2019년의 최대 대기시간을 갖는 지역은 송파구(190분), 경기도 의정부시(189분), 서울시 광진구(163분), 서울시 은평구(149분), 서울시 영등포구(147분)로 나타났다. 2020년의 최대 대기시간을 갖는 지역은 서울시 강북구(492분), 구로구(219분), 동작구(181분), 노원구(171분), 성북구(166분)로 나타났다. 대기시간이 492분인 경우는 출발지가 강북구 수유 제3동이며, 목적지는 강북구 인수동이었다. 배차일시는 20.03.07 오전 7시 4분이었지만 승차일시는 같은 날 오후 3시 17분이었다. 2021년의 최대 대기시간을 갖는 지역은 서울시 송파구(206분), 마포구(186분), 서초구(178분), 양천구(175분), 경기도 남양주시(174분)로 나타났다. 2022년의 최대 대기시간을 갖는 지역은 서울시 영등포구(219분), 강서구(178분), 노원구(172분), 강남구(170분), 관악구(169분)로 나타났다.

〈표 1〉 연도별 최대 대기시간(분) 상위 5개 지역

연도	지역	최대 대기시간(분)
2019	서울시 송파구	190
	경기도 의정부시	189
	서울시 광진구	163
	서울시 은평구	149
	서울시 영등포구	147
2020	서울시 강북구	492
	서울시 구로구	219
	서울시 동작구	181
	서울시 노원구	171
	서울시 성북구	166
2021	서울시 송파구	206
	서울시 마포구	186
	서울시 서초구	178
	서울시 양천구	175
	경기도 남양주시	174
2022	서울시 영등포구	219
	서울시 강서구	178
	서울시 노원구	172
	서울시 강남구	170
	서울시 관악구	169



〈그림 5〉 2019~2022년 수도권 시군구별 최대 대기시간

3) 파생 변수 생성

모델은 개발 환경의 한계로 인해 2021년 1년간의 데이터만을 학습하였다. 모델 구현에 투입할 변수는 TGNNet 알고리즘 선행연구(이도엽 외, 2018)에서 사용한 변수를 차용하였으며 다음과 같다. 차량타입은 ‘대형승용’, ‘중형승합’, ‘중형승용’ 총 3가지 값이 존재한다. 장애유형에 따라 소지하고 탑승할 수 있는 보조기기 크기가 다르므로 배차될 수 있는 차량에 영향을 미치고 이는 곧 대기시간과 연관이 있을 것이라 판단하여 가변수로 생성하였다. ‘승차일시’를 기준으로 요일(7), 공휴일 여부(1), 공휴일 다음날 여부(1) 변수를 생성하였다. 요일은 0~6의 값을 가지며, 월요일은 0, 일요일은 6으로 값을 지정하였다. 2021년의 공휴일 정보는 공공데이터포털 한국천문연구원 제공하러 국경일, 공휴일, 대체공휴일 등을 조회할 수 있는 Open API 데이터를 활용하였다. 이 또한 ‘승차일시’를 기준으로 하였다. 2021년의 공휴일은 총 18일이었으며, 새해(1/1),

설날(2/11~2/13), 삼일절(3/1), 어린이날(5/5), 부처님 오신 날(5/19), 현충일(6/6), 광복절(8/15), 대체공휴일(8/16), 추석(9/20~9/22), 개천절(10/3), 대체공휴일(10/4), 한글날(10/9), 대체공휴일(10/11), 크리스마스(12/25)가 있었다. Open API의 공휴일 정보를 활용해 ‘승차일시’에서 날짜 정보만 추출하여 특정 승차일시가 공휴일이면 1, 아니면 0으로 값을 부여하였다. 공휴일 다음 날에 대한 값은 공휴일 다음 날이면 1, 아니면 0 값을 부여하였다. 또한, 출발지와 목적지 간의 이동거리와 이동시간을 추산한 변수를 생성하였다. 출발지와 목적지의 행정동 단위까지만 제공되는 데이터의 한계로 행정동의 주민복지센터(행정복지센터, 주민지원센터 등)를 기준으로 하여 차로 이동하는 거리와 이동 시간을 계산하였다. 이 과정에서 카카오 맵(Kakao Map)을 웹 크롤링하여 자동차 길 찾기의 이동 거리와 이동시간을 추출하였다. 또한 동일한 행정동 내에서 콜택시를 이용한 경우 측정이 불가능하므로, 임의로 이동거리는 3km, 이동 시간은 5분으로 설정하였다. 따라서 모델에 들어가는 변수는 <표 2>와 같다.

<표 2> 모델 투입 변수 목록

번호	변수 목록	속성
1	차량타입	대형승용, 중형승합, 중형승용
2	승차일시	%Y-%m-%d hh:mm:ss
3	대기시간	초 단위(s)
4	요일	0(월요일) ~ 6(일요일)
5	공휴일	0(공휴일 아님), 1(공휴일)
6	공휴일 다음 날 여부	0(공휴일 다음 날 아님), 1(공휴일 다음 날)
7	이동시간	초 단위(s)
8	이동거리	미터 단위(m)

3. 모델링

1) LSTM 모델 구현

장단기 메모리(Long Short-Term Memory, LSTM) 모델은 시점(time step)이 길어질수록 앞의 정보가 뒤로 충분히 전달되지 못하는 현상이 발생하는 RNN의 한계를 극복한 RNN의 변형 모델이다. LSTM은 은닉층의 메모리 셀에 입력 게이트, 망각 게이트, 출력 게이트를 추가하여 불필요한 기억을 지우고, 기억해야 할 것들을 정한다. 따라서 LSTM은 RNN과 비교하여 긴 시퀀스의 입력을 처리하는 데 탁월한 성능을 보인다. 다양한 분야에서 lstm을 활용한 시계열 예측 연구

가 활발히 진행되고 있다. 그 예로 LSTM은 전력 데이터 예측(이원규, 안소영, 임민섭, 천승태, 2016), 용존산소량 예측(한명수, 박성은, 최영진, 김영민, 황재동, 2020), 온실 내 온도 예측(이정규, 강동현, 이동훈, 2019), 자동차 AS 센터 수리 부품 수요 예측(정동균, 이종화, 이현규, 2022) 등에서 높은 성능을 보인 것으로 드러났다.

LSTM의 수식(이우창, 김양석, 김정민, 이충권, 2020)은 아래와 같다. 은닉층(Forget gate)에서는 유지할 정보를 결정하게 된다. 식 (1)에서는 은닉층에서 이전 셀 상태와 현재 입력 값을 확인하여 Sigmoid layer를 거쳐 어떤 정보를 취하고 버릴지 결정한다. σ 는 활성화 함수, W 는 가중치, b 는 바이어스를 의미한다.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_t) \quad (1)$$

입력층(Input gate)은 앞으로 들어오는 새로운 정보 중 어떤 것을 cell state에 저장할 것인지를 결정한다. 먼저 Sigmoid layer를 거쳐 어떤 값을 업데이트할 것인지 정한 후 tanh layer에서 새로운 후보 vector를 만든다. 식 (2), (3)은 입력층에서 새로운 입력 값을 만드는 수식이다. 수식 (4)는 이전 gate에서 버릴 정보들과 업데이트할 정보를 정했다면, Cell state update 과정에서 업데이트를 진행하는 과정이다.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_t) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = (f_t \times C_{t-1}) + (i_t \times \tilde{C}_t) \quad (4)$$

마지막으로 출력층(Output gate)에서는 어떤 정보를 output으로 내보낼지 결정한다. 식 (5), (6)은 먼저 Sigmoid layer에 input data를 넣어 output 정보를 정한 후 Cell state를 tanh layer에 넣어 Sigmoid layer의 output과 곱하여 output으로 내보내는 과정을 뜻한다.

$$\sigma_t = \sigma(W_\sigma[h_{t-1}, x_t] + b_\sigma) \quad (5)$$

$$h_t = \sigma_t \times \tanh(C_t) \quad (6)$$

2) 정확도 평가

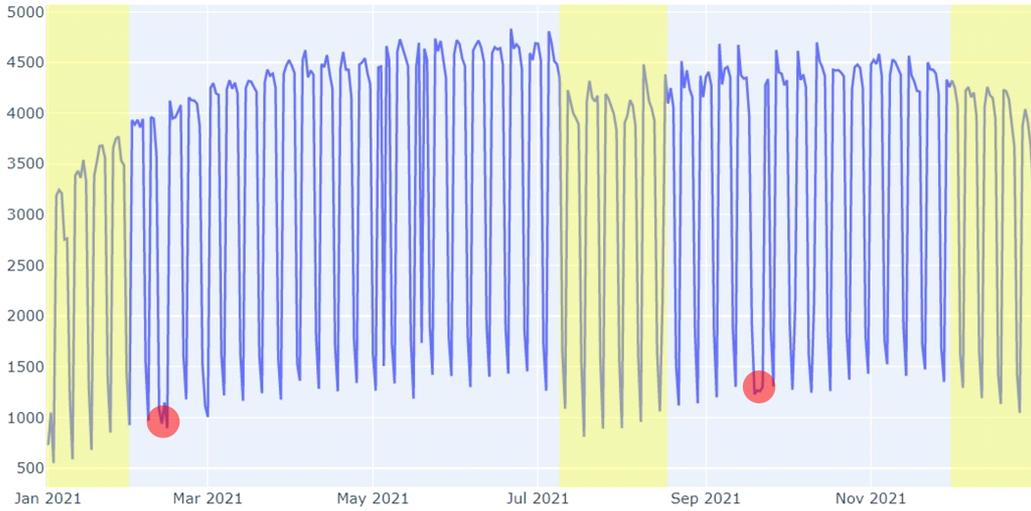
본 연구에서는 예측 모델의 성능을 비교하기 위해 예측 모델의 평가 지표로 많이 사용되는 MSE(Mean Square Error)와 MAE(Mean Absolute Error)를 사용하였다. MSE는 관측 값과 예측 값 사이의 오차의 제곱 값의 평균을 나타낸 지표이다. MAE는 오차의 절댓값의 평균을 뜻한다. 두 평가 지표 모두 낮을수록 예측력이 좋은 모델임을 증명한다. 수식은 (7), (8)과 같다.

$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n} \quad (7)$$

$$MAE = \frac{\sum_{t=1}^n |A_t - F_t|}{n} \quad (8)$$

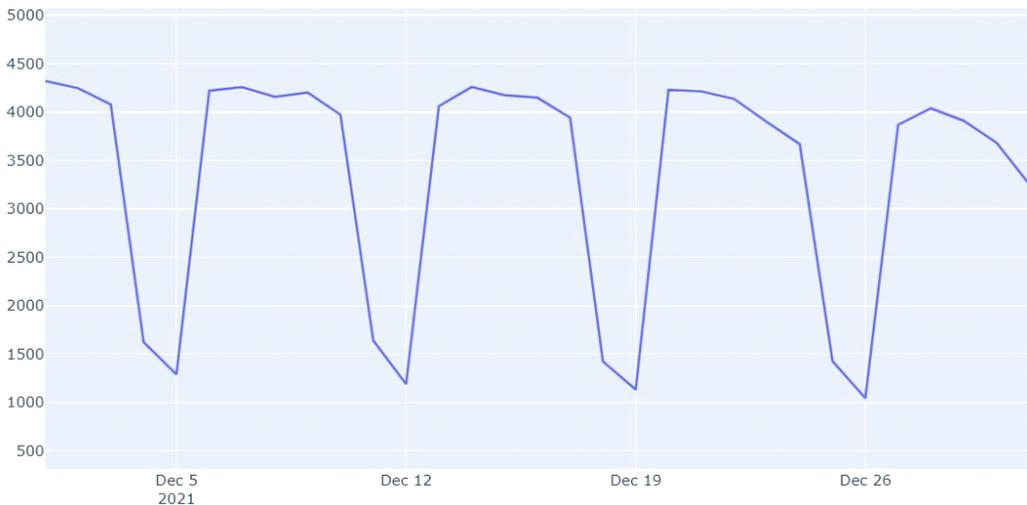
IV. 연구 결과

2021년의 서울시 장애인 콜택시 일별 배차량 그래프는 <그림 6>과 같다. 가장 먼저 계절성을 확인할 수 있다. 여름(8월)과 겨울(1, 12월)에는 다른 기간에 비해 배차량이 현저히 떨어진다. 상대적으로 혹서기, 혹한기에 이동이 줄어드는 특징을 파악할 수 있었다. 또한 빨간색으로 표시된 부분은 평일임에도 낮은 배차량을 보여 상세히 확인한 결과, 명절 연휴인 구정과 추석으로 파악되었다.



〈그림 6〉 2021년 서울시 장애인 콜택시 일별 배차량 연간 주기성

〈그림 7〉은 주간 주기성을 확인할 수 있으며, 예시로 2021년 12월의 그래프를 나열하였다. 5일(일), 12일(일), 19일(일), 26일(일)을 기점으로 1주일 간격으로 배차량이 규칙적으로 변하고, 대체로 평일은 주말보다 배차량이 약 3배정도 많았다.



〈그림 7〉 2021년 서울시 장애인 콜택시 일별 배차량 주간 주기성

〈그림 8〉, 〈그림 9〉는 요일별 변화량을 확인할 수 있는 그래프이다. 2021년 12월 27일, 28일, 29일은 각각 월요일, 화요일 그리고 수요일로 해당 요일이 다른 요일보다 배차량이 가장 높다는 것을 알 수 있다. 다른 기간에도 동일하게 평일에 배차량이 많은 현상을 보였다.



〈그림 8〉 2021년 서울시 장애인 콜택시 일별 배차량 일별 주기성(1)



〈그림 9〉 2021년 서울시 장애인 콜택시 일별 배차량 일별 주기성(2)

〈그림 10〉, 〈그림 11〉은 설날과 추석의 배차량 그래프이며 명절 기간에는 주말처럼 배차량이 크게 떨어지는 것으로 나타났다.



〈그림 10〉 2021년 서울시 장애인 콜택시 설날 배차량



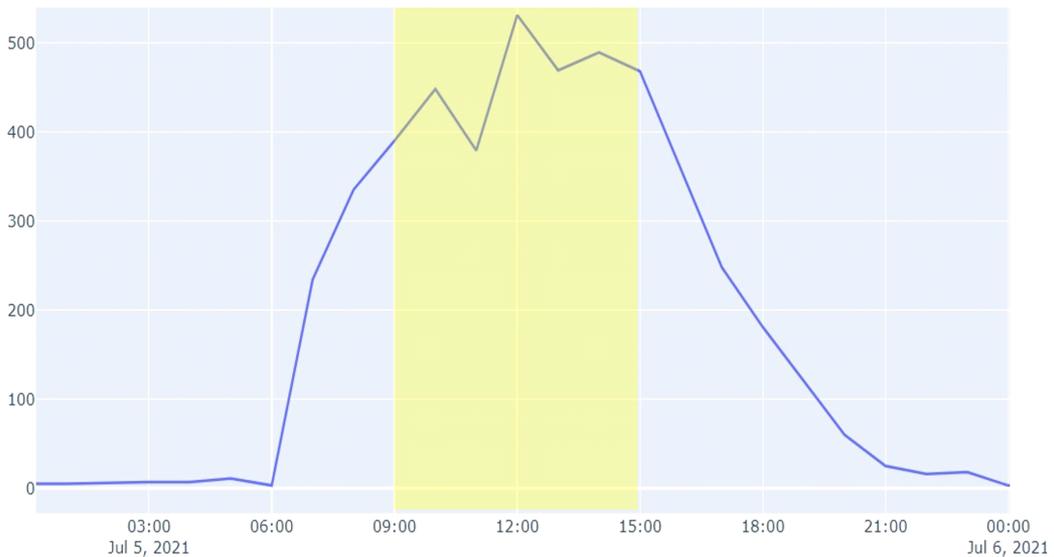
〈그림 11〉 2021년 서울시 장애인 콜택시 추석 배차량

〈그림 12〉, 〈그림 13〉은 하루 동안의 승차시간에 따른 배차량을 알 수 있는 그래프이다. 다음

과 같이 오전 9시부터 오후 3시에 배차 수요가 가장 높음을 알 수 있다. 이는 다른 기간에도 동일한 현상을 보였다.



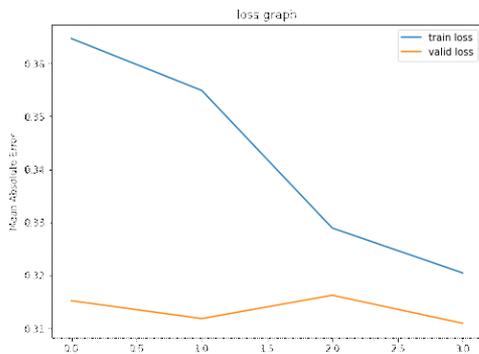
〈그림 12〉 2021년 서울시 장애인 콜택시 시간별 배차량(1)



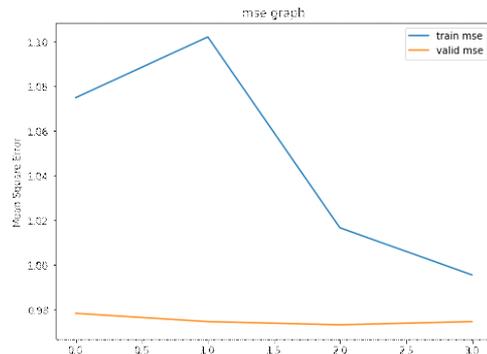
〈그림 13〉 2021년 서울시 장애인 콜택시 시간별 배차량(2)

파악된 현황 특성을 바탕으로 모델 구현을 위해 각 변수를 스케일링하는 작업을 진행하였다. 차량타입, 승차일시, 공휴일 여부, 공휴일 다음 날 여부는 객체(object)로 변환했으며, Python pandas의 get_dummies() 함수를 이용해 더미변수로 치환하였다. 이동거리와 이동시간은 실수(float)로 변환하였다. 또한 전처리된 변수 값들을 모델에 넣기 위해 배열 구조로 변경하였다.

전체 데이터의 70%를 학습 데이터 셋, 30%를 검증 데이터 셋으로 선정하여 모델 학습을 진행하였다. 또한 학습 시 입력되는 데이터 윈도우가 길 경우, 이를 모두 학습하는 것은 많은 비용과 시간이 소모되며 과적합(Overfitting)의 위험이 발생할 수 있다. 이를 방지하고 효율성을 위해 'step'이라는 파라미터를 통해 특정 구간의 데이터만 학습하도록 추출하였다. 본 연구에서는 data_window를 180개, 예측 값을 18개로 설정하였다. data_window가 180개라는 것은 하나의 데이터 윈도우에 180개의 데이터가 입력되는 것을 의미한다. 또한 step을 6으로 설정하였는데, 이를 통해 180개 중 6개마다 데이터를 추출하게 되어 총 30개의 데이터만 추출되게 된다. LSTM 셀의 hidden layer를 32로 설정하고 출력값은 1개이므로 Dense layer에 1을 설정하였다. 모델 컴파일의 최적화 함수는 RMSprop, 손실 함수는 MAE, 측정치는 MSE로 설정하였으며, 학습 프로세스가 모든 샘플에 대해 한 번 실행되는 것을 1 epoch라고 하며 본 연구에서는 가장 낮은 MSE (0.9955)을 보인 4로 에포크를 설정하였다. <그림 14>과 <그림 15>을 통해서 확인할 수 있으며, Train set과 Validation set의 MAE와 MSE가 에포크 4에서 가장 적은 차이를 보였다.



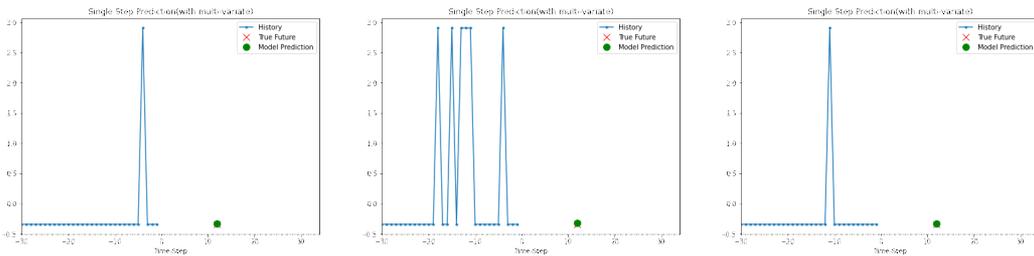
<그림 14> Train set, Validation set의 MAE loss graph



<그림 15> Train set, Validation set의 MSE graph

다음 <그림 16>은 2021년 데이터로 학습된 모델에 랜덤으로 3개의 날짜를 선택하여 해당 날짜의 실제 대기시간과 모델이 예측한 대기시간을 그래프로 나타낸 것이다. 파란 추세선은 학습

데이터로 이루어진 과거 대기시간을 나타낸 History 선이다. 빨간색 표시는 해당 날짜의 대기시간 실제 값, 초록색 점은 해당 날짜에서 모델이 예측한 대기시간 값이다. 두 표시가 가까울수록 정확히 예측되었다고 볼 수 있다.



〈그림 16〉 실제 대기시간과 예측 대기시간 비교

V. 논의

2021년의 장애인 콜택시 일별 배차량을 확인한 결과, 특정 요일에만 배차량 증가, 계절 및 기온에 영향을 미치는 배차량, 추석/설날 같은 명절 기간에는 배차량이 감소하는 것처럼 주기성을 확인할 수 있었다. 그리고 오전 9시에서 오후 3시에는 배차량이 가장 많았고, 월요일에서 수요일이 다른 요일보다 배차량이 높았다. 이를 통해 수요가 유동적인데도 불구하고 공급이 유동적이지 않아서 배차에 문제가 생기고 대기시간이 불필요하게 증가하는 현상이 발생하고 있다고 사료된다. 이를 해결하기 위해 배차량이 높은 시간대 및 기간을 고려하여 그에 맞게 차량과 운전자를 공급하면, 차량이 부족한 문제나 대기시간이 증가하는 문제를 방지할 수 있을 것이다. 예를 들어, 운전자 교대 스케줄을 재조정하여 차량의 운행시간이 골고루 분포되어 원활한 배차 시스템이 가능할 것이다. 실제 운영 대수가 법정 대수를 만족하지 못하는 지역이 태반이므로 지역별 장애인 수를 고려하여 법정 대수를 채우는 것이 시급한 상황이다. 법정 대수에 준하는 장애인 콜택시를 갖춰야 하는 것이 최우선적이다. 또한 일반 택시 운전자들을 프리랜서 형태로 일시적으로 공급하여 공급난을 해결할 수 있다. 택시 운전자는 운전 경력이 풍부하고 무사고 운전 경력을 갖춘 인력으로, 이들에게 적절한 인센티브와 장애 및 장애인 서비스 관련 교육을 충분히 지원한다면 택시업과 병행이 가능할 것으로 판단된다. 예를 들어, 장비 작동 방법, 돌발 상황 시 대처법, 차량 안전 기준 등을 표준화 및 의무화할 필요가 있다. 부수적으로 일반 택시 운전자들도 장애인의 이동권에 대한 이해를 높여 장애인 이동권을 보장하기 위해 함께 노력해나가는 효과를 얻을 수

있다. 장애인의 이동권은 결코, 장애인 및 보호자 또는 관련 기관 종사자뿐만 아니라 사회 모두가 점진적으로 해결해 나가야 할 문제이기 때문이다. 지역별, 혹은 장애 인구 밀집 지역을 우선으로 차고지를 확대한다면 특정 지점에 차량이 치우치는 현상을 줄여 대기시간의 편차를 좁히는 데에 도움이 될 것이라 사료된다. 또한 서울시 및 수도권 이외의 타 지역에서도 장애인 콜택시와 같은 특별교통수단 현황을 파악할 수 있는 데이터를 구축할 필요가 있다. 공공 데이터가 공개되어 놓여촌처럼 장애인 이동 사각지대에 놓여있는 도 단위 지역에 대한 연구도 활발히 진행된다면 지역별 양극화가 심화되지 않도록 개선해 나갈 수 있을 것이다.

여전히 장애인은 비장애인 대비 적게는 4배, 많게는 6배까지 이동을 위해 대기해야 한다. 전 장면은 현재도 장애인 이동권 보장을 위한 시위를 2021년 12월 3일부터 현재까지 지속하고 있다. 본 연구를 통해 아직도 장애인의 이동권은 비장애인과 비교조차 불가능하게 부족한 수준인 것을 확인하였다. 장애인의 이동권을 보장하기 위해서는 지금보다 훨씬 많은 관심이 필요하고 사회 구성원 모두가 함께 바뀌 나가야 할 부분이 존재한다. 이러한 측면에서 장애인의 원활한 이동을 위하여 일반 택시들을 프리랜서 형태로 배차 수요가 높은 시간, 요일, 월에 적절히 배치하여 불균형적인 수요에 대비하는 대안을 제시한다. 더불어, 택시 운전자들이 최근 토로하는 승차난을 일부 해소할 수 있어 상생하는 방향으로 발전할 수 있다.

본 연구가 갖는 기존 연구와의 차별성은 다음과 같다. 첫째, 2019.01.07 ~ 22.06.30 기간 동안에 작성된 실제 장애인 콜택시 이용 후기글을 웹 크롤링 및 시각화를 하였고, 이들이 장애인 콜택시에 대해 어떠한 인식과 불편을 가지고 있는지 파악하여 배차와 대기시간이라는 구체적인 문제점을 발견하였다. 그리고 이를 개선하고자 장애인 콜택시 운영 현황 Open API 정형 데이터를 분석하였다. 둘째, 기존 콜택시 데이터에 존재하는 변수 이외에 모델 구현과 모델의 성능을 높이기 위해 Feature Engineering을 진행하였다. 승차시간과 배차시간의 차이를 이용하여 대기시간을 구하였고, 승차시간을 기준으로 승차한 날의 요일, 공휴일 여부, 공휴일 다음 날 여부를 구하였다. 공휴일의 경우, Python에서 제공하는 모듈 및 패키지는 부정확한 정보가 있다는 한계를 발견하여 공공데이터포털의 한국천문연구원에 제공하는 정보를 이용하여 데이터의 정확도를 높이고자 하였다. 카카오 맵의 길 찾기 서비스에서 제공하는 출발지와 목적지 간의 이동거리와 이동시간 정보를 웹 크롤링으로 생성하였다. 동일한 행정동 내에서 콜택시를 이용한 경우 측정이 불가능하므로, 임의로 이동거리는 3km, 이동 시간은 5분으로 설정하였다. 셋째, LSTM 모델의 강점을 활용하였다. 시점이 길어질수록 앞의 정보가 뒤까지 전달하지 못하는 현상이 발생하는 RNN의 한계를 극복하고자 하였다. 또한 LSTM은 은닉층의 메모리 셀에 입력 게이트, 망각 게이트, 출력 게이트를 추가하여 기억해야 할 것만 정한다. 따라서 과거의 정보를 훨씬 잘 반영할 수 있다.

VI. 결론

본 연구에서는 장애인들이 일상생활에서 가장 불편함을 느끼는 이동 문제를 지적하였고, 이를 해결하기 위한 대기시간 예측을 시도하였다. 따라서 연구가 가지는 의의는 다음으로 요약할 수 있다. 텍스트 마이닝 기법을 통해 장애인 콜택시 이용자들이 현재 어떠한 불편 사항을 가지고 있는지 구체적으로 파악하였다. 또한, 이를 해결하기 위해 장애인 콜택시 Open API 정형 데이터를 활용하여 대기시간 예측 모델을 구성하였다. 주어진 데이터를 최대한 활용하여 모델에 사용할 파생 변수를 생성하였다.

하지만 본 연구는 아래와 같은 한계점이 존재한다. 첫째, 사용된 콜택시 데이터는 출발지와 목적지가 동 단위로 공개되었고, 출발지와 목적지 정보가 변경된 지역명이나 통합된 지역 명을 반영하지 못하는 경우가 존재하였다. 따라서 동별 주민 센터나 행정복지센터 등의 장소로 특정하여 분석을 진행하였다. 정확한 위·경도를 사용하지 못하였기에 데이터의 정확도가 떨어질 수 있다. 또한, 동일한 동 내에서는 이동거리를 측정하지 못하여 임의로 5분과 3km로 설정하였다. 이는 서울특별시 콜택시의 정확한 데이터를 확보할 시 해결될 수 있다고 판단된다.

둘째, 각 지역의 교통량과 그에 따른 혼잡도를 고려하지 못하였다. 교통량에 따른 교통체증 문제는 대기시간과 밀접한 관계가 있으므로 추후 연구에서는 혼잡도를 고려하여 모델의 정확도를 높일 수 있을 것이다. 이는 카카오 맵의 기능 중 '미래 운행 정보'를 이용하여 승차 시간대의 이동 시간과 이동 거리를 스크래핑 한다면 더 정확한 결과를 도출함으로써 해결할 수 있다. 하지만 본 연구에서는 편의상 출발지와 목적지가 동일하다면 동일한 이동 시간과 이동 거리를 할당하여 진행하였다.

셋째, 개발 환경의 제한으로 다양한 하이퍼 파라미터 튜닝을 시도하지 못하였다. 추후 연구에서는 원활한 개발 환경을 마련하여 모델의 정확도를 높일 수 있을 것으로 사료된다.

본 연구와 함께 장애인과 비장애인의 이동권을 동일한 수준까지 끌어올릴 수 있는 연구가 꾸준히 진행되어야 할 것이다. 장애인 이동권 문제는 이전부터 계속 수면 위로 올라온 문제이지만 여전히 사회 곳곳에서 개선을 필요로 하며, 사회 전체가 더 많은 관심을 가져야 한다는 것을 시사한다. 구체적으로 비장애인과 마찬가지로 이동권에 제약을 받지 않도록 장애인의 콜택시 대기 시간을 줄일 수 있는 연구가 지속적으로 필요하다. 본 논문이 그 목적에 다가설 수 있는 하나의 기초 자료로 활용되어 도움이 되기를 기대한다.

참고문헌

- 김건욱, 윤대식, 김종진. (2020). 빅데이터를 활용한 교통약자 특별교통수단 통행수요 분석; 대구광역시를 사례로. 대구경북연구, 19(2), 43-61.
- 김유민. (2022.09.11.). 택시기사들 “통닭 배달도 5000원인데...사람은 3800원”. 서울신문.
<https://n.news.naver.com/mnews/article/081/0003300823?sid=102>
- 공공데이터포털. 전국교통약자이동지원센터정보표준데이터.
<https://www.data.go.kr/data/15028207/standard.do>
- 공공데이터포털. 한국천문연구원_특일 정보.
<https://www.data.go.kr/data/15012690/openapi.do>
- 국가통계포털. 시군구별,장애정도별,성별 등록장애인수.
https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_11761_N009&conn_path=I2
- 딥 러닝을 이용한 자연어 처리 입문. 장단기 메모리(Long Short-Term Memory, LSTM).
<https://wikidocs.net/22888>
- 박성제. (22.03.31). “[팩트체크] 콜택시 타려고 2시간 기다리는 경우 많다?”. 연합뉴스.
<https://www.yna.co.kr/view/AKR20220331051400502>
- 박진혁, 장영재. (2015). 서울시 장애인 콜택시 고객 대기시간 감소를 위한 자동배차 알고리즘 및 최적 차량 공급 제안. 한국경영과학회, 2015(4), 825-847.
- 박형수, 김훈기, 김영성. (2017). 교통약자를 위한 시간 복잡도를 반영한 자동배차 알고리즘. 전자공학회, 154(6), 43-52.
- 서울 열린데이터광장. 장애인콜시스템.
<https://data.seoul.go.kr/dataList/OA-15558/A/1/datasetView.do>
- 석종수. (2014). 인천광역시 장애인콜택시 운행 특성 분석. 교통기술과정책, 11(3), 51-58.
- 성광환, 신유열, 윤운진. (2022). 지역별 장애인 콜택시의 접근 형평성 분석. 한국경영과학회 학술대회논문집, 1796-1807.
- 손종훈, 김도경. (2021). 생존분석을 활용한 장애인콜택시 이용자 대기시간 영향요인 연구. 대한교통학회, 496-497.
- 신용은, 최혜미, 송기욱, 이희대. (2014). 특별교통수단 이용자 통행패턴 분석 - 경상남도 사례 -.

대한토목학회논문집, 34(1), 213-221.

이민경. (2018). 장애인의 사회활동 및 문화·여가활동 실태와 정책과제. 한국보건사회연구원, 263, 73-88.

이병화, 양희택. (2017). 서울과 경기도의 장애인 콜택시 이용현황 빅데이터 분석 연구. 예술인문 사회융합멀티미디어논문지, 7, 925-934.

이우창, 김양석, 김정민, 이충권. (2020). 머신러닝을 이용한 철광석 가격 예측에 대한 연구. 한국 산업정보학회논문지, 25(2), 57-72.

이원규, 안소영, 임민섭, 천승태. (2016). LSTM을 이용한 전력 데이터 예측. 한국정보과학회 학술발표논문집, 693-695.

이재길. (2020). 모델리티 AI 동향. 한국과학기술원 정보과학회지, 38(8), 42-48.

정동균, 이종화, 이현규. (2022). LSTM 인공지능망을 이용한 자동차 AS 센터 수리 부품 수요 예측 모델. 한국경영정보학회 2022 경영정보관련학회 춘계통합학술대회, 439-443.

정진형, 박호성, 여윤구, 김학빈, 강준혁. (2018). Queueing System을 이용한 교통약자 특별교통수단의 효율적 증차 방안. 대한산업공학회, 2837-2851.

조태호. (2022). 모두의 딥러닝 개정 3판. 서울: 길벗 출판사, 140.

조현기. (2022.07.25). 전장연 “서울경찰청장 공식 사과해야 수사받겠다”. 뉴스1. <https://n.news.naver.com/mnews/article/421/0006238230?sid=102>

한명수, 박성은, 최영진, 김영민, 황재동. (2020). 시계열 분석을 이용한 진동만의 용존산소량 예측. 해양환경안전학회지, 26(4), 382-391.

한연준, 노은정, 조범래. (2022). 랜덤포레스트와 순환신경망 기반 수소 충전소 혼잡도 예측 모델에 관한 연구. 한국자동차공학회 춘계학술대회, 2022(6), 1051-1060.

홍두선, 가은한, 하성현, 이청원. (2018). 서울시 장애인콜택시 대기시간 예측 LSTM 모델의 적정 하이퍼파라미터 선정. 대한교통학회, 78, 449-454.

人CoDOM. 2021. LSTM. <http://www.incodom.kr/LSTM>

Feeley, C., Spasovic, L. N., Dimitrijevic, B., & Lee, J. (2016). Develop and Pilot Test Smart Phone /Tablet App for Paratransit Demand-Response Passenger Pick-up Alerts to Assist Passengers with Disabilities and Reduce No-Shows and Dwell Times, Rutgers University. Center for Advanced Infrastructure &

Transportation, 1-14.

Garnier, C., Trepanier, M., & Morency, C. (2020). Adjusting dwell time for paratransit services, Transportation Research Record, 2674(9), 638-648.

Lee, D., Jung, S., Cheon, Y., Kim, D., & You, S. (2018). Forecasting taxi demands with fully convolutional networks and temporal guided embedding.

원고접수 : 2022. 9. 18.

수정원고접수 : 2022. 11. 8.

게재확정 : 2022. 11. 28.

Abstract

A Study for Estimating Waiting Time of Taxi Services for the Disabled People based on LSTM

Lee Jeong Min*·Park Yeong Bin**

Currently, the problem of the mobility rights of the disabled is still unresolved. First of all, we performed Word Cloud and Semantic Network Analysis(SNA) of the reviews of call taxi users with disabilities. The results showed there were inconveniences about the waiting time of the call taxi for the disabled and the traffic allocation system. As the years go by, the waiting time for a call taxi tends to decrease, but on average, it still takes about 30 minutes up to 2 hours or more to get the taxi. In addition, in most regions, the number of call taxis did not satisfy the number of legal taxi. In this study, by collecting and analyzing public data, we confirmed the actual conditions of the service operation fall short of expectations and call taxi users with disabilities felt the inconvenience. Therefore, we propose an approach to train Long-Short-Term Memory (LSTM) architecture in order to predict a sequence of next waiting time. The proposed LSTM model achieved performance in the testing data set indicated by the MSE(Mean Square Error) values of 1.0749, 1.1021, 1.0166 and 0.9955 at epoch 4 and we finally chose the last model. We come up with a plan to reduce the waiting time of call taxis for the disabled from the result. Furthermore, we look forward to improving the quality of life of the disabled and ultimately guaranteeing the mobility rights of the disabled in our society.

Keywords : the mobility rights of the disabled, call taxi for the disabled, text mining, prediction of waiting time, LSTM

* Korean National Police University

** Department of Information Statistics, Gangneung-won National University