

소상공인들을 위한 딥러닝 기반의 상권 방문자 수 예측 시스템

황병일, 여동훈, 김다현, 이수민, 김연이, 서희주, 최상민, 서영주*

포항공과대학교 인공지능연구원, *포항공과대학교 인공지능연구원

{bihwang, tgc01098, kdhyun8011, hibokchi, ouiouio1131, lilysunday, bestmin, yjsuh*}@postech.ac.kr

Deep Learning-Based Visitor Number Prediction System for Small Business Owners in Commercial Districts

Hwang Byeong Il, Yeo Dong Hun, Kim Dahyun, Lee Soo Min, Kim Yeon Li, Seo Hee Ju,

Choi Sang Min, Suh Young Joo*

POSTECH Institute of Artificial Intelligence

요약

음식점 등의 상점을 운영하는 소상공인들은 상점이 위치한 상권의 유동 인구에 크게 의존하고 있다. 유동 인구의 변동성은 매출에 큰 영향을 미치며, 재료 준비량 등 운영 전략에도 중요한 역할을 한다. 그러나, 날씨, 행사 등 외부 요인에 의해 유동 인구는 크게 변동하며, 이를 예측하는 것은 매우 어려운 일이다. 이런 문제를 해결하기 위해 본 논문에서는 캡티브포털을 활용한 플랫폼을 구축하였고, 이 플랫폼은 소상공인들의 마케팅 전략 수립 및 가게 운영을 지원할 수 있도록 설계되었다. 또한, 이 플랫폼에는 유동 인구 센서가 포함되어 있어, 실시간 상권 방문자 수를 추적하고 예측할 수 있다. 본 논문에서는 경북 구미시 진평동의 14개 상점에 유동 인구 센서를 설치하고 3개월 동안의 데이터를 수집하였다. 이 데이터를 바탕으로, 딥러닝 기반의 상권 방문자 수 예측 시스템을 개발하였다. 시계열 데이터 분석에 강한 1D-CNN과 LSTM 단일 모델, 그리고 두 구조를 결합한 모델의 예측 성능을 비교 분석하였다. 분석 결과, 1D-CNN 단일 모델이 MAE(Mean Absolute Error) 7.55로 가장 우수한 예측 성능을 보였다. 이 연구를 통해, 소상공인들은 실시간으로 정확한 방문자 수를 예측하고 최적의 상점 운영 전략을 수립할 수 있게 되었다. 이는 소상공인들의 경제적 손실을 최소화하는데 큰 도움이 될 것으로 기대된다.

I. 서론

소상공인들, 특히 음식점을 운영하는 사업자들은 그들의 사업체가 위치한 상권의 유동 인구에 크게 의존하고 있다. 상권의 유동 인구는 음식점의 매출액, 재료 준비량, 그리고 이벤트 계획 등 가게 운영의 핵심적인 모든 측면에 중대한 영향을 미치는 중요한 요소이다[1].

유동 인구를 정확히 예측하면, 소상공인들은 매출액을 증가시키고, 비용을 줄이며, 고객 만족도를 향상시킬 수 있는 전략을 세울 수 있다. 예를 들어, 유동 인구가 많을 것으로 예측되는 날에는 더 많은 재료를 준비하고, 프로모션 이벤트를 계획할 수 있으며, 반대로 유동 인구가 적을 것으로 예측되는 날에는 재료 손실을 최소화하고, 필요한 직원 수를 줄여 운영 비용을 절감할 수 있다.

그러나, 유동 인구는 갑작스러운 날씨 변화나 지역 내에서 발생하는 행사 등의 외부적인 요인에 의해 큰 변동성을 보이기 때문에 이를 예측하는 것은 쉽지 않다. 이러한 문제를 해결하기 위해 본 논문에서는 실시간으로 유동 인구를 예측해 소상공인들이 유동적이고 효율적인 가게 운영을 가능하게 하는 딥러닝 기반의 상권 방문자 수 예측 시스템을 제안하고자 한다.

이를 위해 우선, 마케팅 지원과 방문자들의 가게 방문 시 주문 기능 등 다양한 편의 기능이 탑재된 캡티브포털을 활용한 플랫폼을 구축하였다 [2][3]. 그리고 이 플랫폼에 통합될 상권 방문자 수 예측 모델을 개발하기 위해, 경북 구미시 진평동에 위치한 14개의 상점에 유동 인구 센서를 설치하고 3개월 간의 데이터를 수집하였다. 수집된 데이터를 바탕으로, 시계열 데이터 분석에 강한 1D-CNN과 LSTM 단일 모델, 그리고 두 구조를 결합한 모델의 성능을 비교 분석하였다. 이 중 1D-CNN 단일 모델

이 MAE 7.55로 가장 우수한 방문자 수 예측 성능을 보였으며, 이를 활용하여 예측 시스템을 구축하였다.

II. 관련 연구

1. 유동인구 빅데이터 딥러닝을 활용한 상가 권리금 적절성 검토 시스템

이동통신기지국 신호 기반으로 수집된 유동 인구 데이터와 특정 지점에서 수집된 유동 인구 공공 데이터를 결합하여 유동인구 빅데이터를 구축하였다[4]. 이 데이터를 활용하여 딥러닝 기반으로 유동 인구의 변화 추이를 분석하는 연구가 진행되었다. 이 연구는 특정 시점부터 현재까지의 유동 인구 빅데이터 변화를 분석하여 기대 손님수의 변화 추이를 도출하였고, 이를 통해 상가 권리금에 대한 적절성을 검토하는 데 지원하였다.

그러나 해당 연구는 하루 전체의 방문자 수를 예측하려면 상당한 데이터 구축 비용이 들어가며, 실시간 예측을 제공하지 못하는 한계를 가지고 있다.

2. 스마트시티 교통을 위한 POI(Point Of Interest)유동 인구 데이터 기반의 방문수요 분석

또 다른 연구에서는 상권 내 POI(Point of Interest)의 일일 인구 유동 패턴을 셀 단위로 분석하여 상업적 측면에서의 방문자 수요를 파악하였다 [5]. 해당 연구에서는 SKT에서 제공한 유동 인구 데이터와 국토교통부의 GIS 데이터, 그리고 소상공인시장진흥공단의 데이터를 활용하여 분석을 진행하였다. 업종별로 유의미한 유동 인구 패턴을 찾아내고, 평일과 주말

의 방문수요 차이를 확인하여 스마트시티 교통 시스템을 설계하였다. 이 연구는 보다 정밀한 유동인구 수에 대한 분석을 통해 요일 별, 시간별 방문 수요 패턴을 파악하였다.

그러나 해당 연구 역시 특정 행사의 유무 등의 외부적 요인을 반영하지 못하고, 실시간 방문자 수 예측을 제공하지 못하는 한계를 가지고 있다.

III. 딥러닝 기반의 상권 방문자 수 예측 시스템

본 논문에서는 캡티브 포털을 활용한 플랫폼 내 설치된 유동인구 센서를 통해 수집된 데이터를 기반으로 실시간 방문자 수 예측하는 시스템을 제안한다.

1. 유동인구 센서 데이터

사용된 유동인구 센서 데이터는 2023년 6월 14일부터 2023년 9월 7일까지 경북 구미시 진평동에 설치된 14개의 센서로부터 수집되었다. 총 13,160개의 데이터는 1시간마다 방문자 수를 누적하여 기록했다. 시계열 데이터의 특성상 적절한 입력 길이를 설정해야 하므로, 전체 데이터 개수와 특성을 고려하여 3일치 데이터를 입력으로 활용했다. 또한, 영업시간 외인 12시간을 제외하고 12시간만을 입력으로 사용하기 위해 36개의 데이터를 입력으로 설정했다.

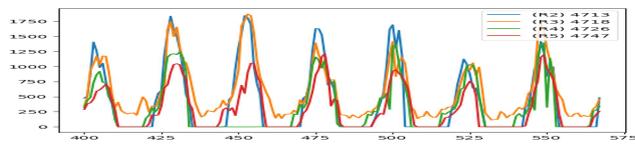


[그림1] 학습/테스트 데이터셋 분할 예시

36시간의 데이터는 [그림 1]과 같이 하루씩 시프팅(shifting)하여 총 7,840개의 데이터를 생성했다. 그리고 이 데이터를 학습, 검증, 테스트 데이터셋으로 8:1:1의 비율로 분할했다.

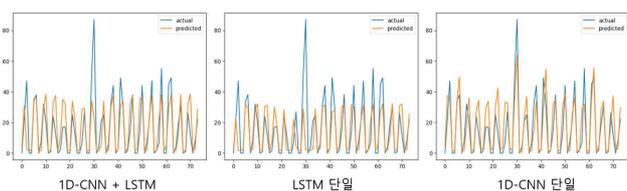
2. 방문자 수 예측 모델 분석 및 개발

최적의 방문자 증감 추세 예측 AI 모델을 개발하기 위해 RNN 및 1D-CNN 기반의 알고리즘을 활용했다. 현재 시간을 기준으로 이전의 36시간의 입력 데이터를 사용하며, 이 데이터는 미운영 시간을 제외한 것이다. [그림 2]와 같이 1시간, 2시간, 3시간 뒤의 각 방문자 수를 예측하였다.



[그림2] 범주 순으로 실제 방문자수, 24시간, 1시간, 3시간 예측 결과

실험을 위해 사용한 모델은 1D-CNN+LSTM, LSTM, 1D-CNN 단일 모델이며, 이 모델들은 실제 방문자 수와 예측 방문자 수의 차이를 최소화하는 방향으로 학습을 진행했다.



[그림3] 모델별 실제 방문자 수와 예측된 방문자수 추이

[표1] 모델 별 성능치(MAE)

Model	MAE
1D-CNN+LSTM	7.74
LSTM 단일	8.07
1D-CNN 단일	7.55

실험 결과, [표1]과 같이 1D-CNN 모델이 MAE 7.55로 가장 우수한 성능을 보였다. 이는 유동인구 센서 데이터의 복잡도가 단순하여 1D-CNN을 통한 방문자 수 예측이 가장 우수한 성능을 보인 것으로 분석된다.

III. 결론

본 연구에서는 캡티브포털을 활용한 플랫폼에 적용 가능한 실시간 방문자 수 예측 시스템을 개발하였다. 이 시스템은 플랫폼에 설치된 유동인구 센서를 통해 수집된 방문자 수 데이터를 기반으로, 현재 플랫폼이 운영 중인 구미 진평동의 상권에 맞게 방문자 수를 예측하도록 설계되었다.

향후에는 본 논문에서 제안한 상권 내 방문자 수 예측 시스템을 발전시켜, 날씨 등의 기상 요건을 고려하도록 할 계획이다. 지속적인 데이터 축적을 통해 연 단위의 방문자 수 데이터를 수집하게 되면, 상권 내에서 규칙적으로 발생하는 행사 일정과 기상 상황 등 외부 요인까지 고려한 방문자 수 예측이 가능해질 것으로 보인다. 이렇게 되면, 소상공인들은 해당 상권에 맞는 마케팅 전략을 수립하거나, 가게 운영에 필요한 자원을 보다 효율적으로 관리하는 등의 작업을 수행할 수 있을 것이다. 이를 통해 소상공인들의 가게 운영을 더욱 효율화하고, 경제적 손실을 최소화하는데 큰 도움이 될 것으로 기대된다.

ACKNOWLEDGMENT

이 논문은 2023년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업입(No.2022R1A6A1A03052954)

참 고 문 헌

- [1] 김현철, 이승일, “서울시 골목상권 매출액에 영향을 미치는 요인에 관한 연구”, 서울시연구 제20권 제 1호, pp.117-134, 2019.03.
- [2] 황병일, 김동주, “방문자 추천을 위한 개인 행동 범주 예측”, 한국 컴퓨터정보학회 하계학술대회 논문집 31권 2호, pp5-6, 2023.07
- [3] 여동훈, 황병일, 김동주, “단어 임베딩 모델 기반 캡티브 포털 메뉴 추천 시스템”, 한국 컴퓨터정보학회 하계학술대회 논문집 31권 2호, pp.11-12, 2023.07
- [4] 김명락, “유동인구 패턴정보 기반의 미래입점객수 예측방법 및 장치”, 대한민국, 1018101690000, 2017.12.12
- [5] 홍윤재, 이기쁨, 이주현, 양효준, 이지현, “스마트시티 교통을 위한 POI(Point Of Interest) 유동인구 데이터 기반의 방문수요 분석”, 2023 한국컴퓨터종합학술대회 논문집, pp.229-231