

딥 러닝을 이용한 다중 도로구간 속도 예측

손지원*, 송준호*, 김남혁**, 김태헌**, 박성환**, 김상욱†*

*한양대학교 컴퓨터 소프트웨어학과

**현대자동차

{tinybeing, hahadori, wook}@hanyang.ac.kr, {namhyuk.kim, Kim.Taeheon, sunghwan.park}@hyundai.com

A Deep Learning Based Traffic Speed Prediction on Multiple-Roads

Jiwon Son*, Junho Song*, Namhyuk Kim**, Taeheon Kim**, Sunghwan Park**, Sang-wook Kim†*

*Dept. of Computer Science, Hanyang University

**Hyundai Motor Company

요 약

최근 활발히 진행되는 교통 속도 예측 연구는 기존에는 하나의 모델로 하나의 도로구간에 대해서만 예측하는 문제를 주로 다루었다. 그러나 하나의 도로구간을 하나의 속도 예측 모델로 예측할 시, 도로구간마다 모델이 존재하여야 하므로 모델의 예측 비용이 도로구간의 수만큼 증가한다. 본 논문에서는 하나의 모델을 통해 다수의 도로구간에 대한 속도를 예측하는 다중 도로구간 속도 예측 모델을 제안한다. 제안하는 다중 도로구간 속도 예측 모델은 기존의 단일 도로구간 속도 예측 모델 대비 정확도를 보존하면서, 그 예측 비용을 크게 감소시켰다.

1. 서론

교통 속도 예측 연구는 과거의 교통 속도를 기반으로 미래 속도를 예측하는 연구이다. 도로구간의 교통 속도는 교통 관리, 최적 경로 탐색 등의 문제에 핵심적인 정보이기 때문에, 교통 속도 예측 문제가 활발히 연구되고 있다 [1].

과거의 교통 속도 예측은 시계열 예측, 통계적 예측, 기계 학습과 같은 방법으로 연구되었다 [2-4]. 최근 연구에서는 데이터의 특성을 모델이 직접 학습하여 복잡한 데이터의 표현이 가능한 딥 러닝 모델이 활용되고 있다 [5]. 또한, 교통 속도만을 이용하던 과거에 비해 속도 외적인 요인 (외부요소)들을 이용하는 연구들이 늘어나는 추세이다. 날씨, 사고 및 시간 등의 외부요소들은 교통속도와 밀접한 연관이 있으며, 여러 연구에서 이러한 외부요소 데이터를 이용하여 예측 정확도를 향상시키고 있다 [6-8].

기존의 교통 속도 예측 연구는 주로 도로구간 당 한개의 예측 모델로 속도를 예측하였다. 그러나, 실용적 측면에서 전국 도로구간 수만큼의 모델을 관리하고 이용하는 것은 막대한 비용이 요구된다.

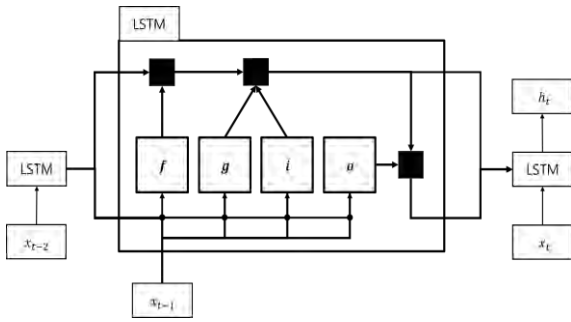
따라서, 이러한 비용 절감을 위해 단일 모델로 다수의 도로구간을 예측하는 모델의 필요성이 커지고 있다 [9]. 그러나 다양한 도로구간 속도 경향을 학습하는 것은 어려운 일이다. 모델이 입력하는 속도를 도로구간에 대한 구분 없이 무작위로 학습할 경우, 각기 다른 도로구간의 속도 경향이 서로 학습에 방해가 된다.

이를 해결하기 위해, 본 논문에서는 다수의 도로구간 속도를 예측하는 하나의 딥 러닝 모델을 제안한다. 다양한 도로구간의 식별을 위하여 도로구간의 고유번호인 링크 아이디를 이용하였다. 예측 모델은 시계열 데이터 예측 시 자주 사용되는 딥 러닝 모델인 Long-Short Term Memory (LSTM, 그림2) [10]을 활용하였다.

제안하는 모델의 검증을 위해 실제 경부고속도로상의 10개의 도로구간을 선택하여 실험한 결과, 10개의 모델을 이용하여 10개의 도로구간의 속도를 예측한 경우와 단일 모델을 이용하여 10개의 도로구간의 속도를 예측한 경우의 정확도의 차이가 거의 없었다. 반면, 10개 도로구간의 속도를 1회씩 예측하는

† 교신 저자

데에 소요되는 시간은 약 82% 단축되었다.



(그림 1) LSTM 구조.

2. 제안하는 방법

다수의 도로구간 속도를 예측하는 1개의 모델을 다중 도로구간 속도 예측 단일 모델 (Single Model for Multi-roads speed Prediction, 이하 SMMP)이라 하고, 1개의 도로구간 속도를 각각 예측하는 1개의 모델은 단일 도로구간 속도 예측 단일 모델 (Single Model for Single-road speed Prediction, 이하 SMSPP)이라 한다. SMMP, SMSPP 모두 구조는 그림 1과 같이 LSTM으로 동일하며 과거 n분의 데이터 및 외부요소 데이터를 입력하여 해당 도로구간의 t분 후의 속도 값을 예측한다.

각 모델에 입력되는 과거 n분의 데이터는 과거 속도 데이터와 외부요소 데이터로 이루어진다. 과거 속도 데이터는 5분 단위의 과거 속도를 최소-최대 정규화 한 벡터이다. 외부요소 데이터는 날씨, 시간, 이벤트로 구성되고 [7]의 구조와 같다.

SMMP의 경우 SMSPP와 달리 다중 도로구간의 과거 데이터 입력을 위해 각 도로구간의 링크 아이디를 추가로 입력한다. 링크 아이디는 도로구간마다 부여되는 고유 번호로 최대 7자리의 숫자로 표현된다. 따라서 각 링크 아이디는 2진수로 변환한 뒤 0과 1로 이루어진 벡터로 입력한다.

3. 실험

3.1 실험 환경

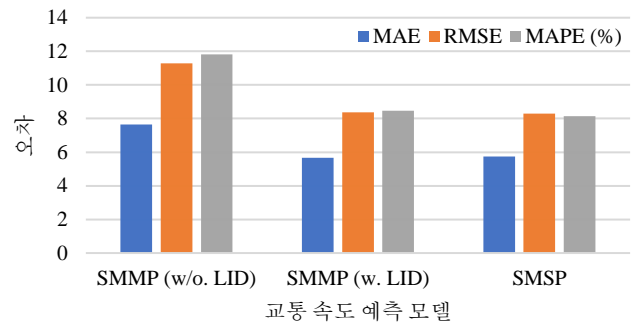
전국 도로구간 중 경부고속도로 상의 10개 도로구간의 2018년 11월부터 2019년 12월 간의 데이터를 실험에 이용한다. 데이터 수집 및 실험 비용의 한계로 10개 도로구간으로 제한하였다. 속도, 이벤트, 링크 아이디 데이터는 현대자동차에서 전달되었고, 날씨 데이터는 기상청의 기상자료개방포털에서 제공받았다. 시간 데이터는 직접 가공하였다.

10개의 도로구간에 대해 모두 학습을 한 SMMP와 각각을 학습한 10개의 SMSPP를 비교한다. 과거 60분 분량의 데이터 입력으로 60분 후의 속도를 예측하였

다. 정확도는 5-cross validation으로 검증하였고 사용된 지표는 평균 절대 오차 (MAE), 평균 제곱근 오차 (RMSE), 평균 절대 백분율 오차 (MAPE)이다.

3.2 실험 1. 전체 예측 정확도 비교

실험 1에서는 링크 아이디를 사용하지 않은 SMMP, 사용한 SMMP, 그리고 SMSPP의 정확도를 비교하였다.



(그림 2) LID 사용 여부에 따른 SMMP 및 SMSPP의 정확도.

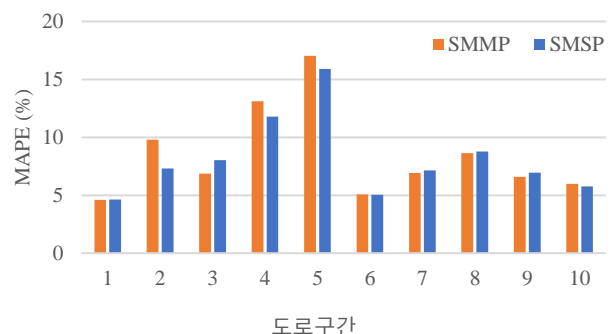
그 결과는 그림 2로 나타냈다. 그림 2에서 'SMMP (w/o. LID)' 는 링크 아이디를 이용하지 않은 SMMP, 'SMMP (w. LID)' 는 이용한 SMMP를 의미한다. 정확도는 10개 링크에 대한 정확도의 평균값이다.

정확도는 SMSPP, SMMP (w. LID), SMMP (w/o. LID) 순서로 높다. MAPE 기준으로 각각 8.13, 8.46, 11.82이다. SMMP (w/o. LID) 의 경우, SMSPP에 비해 오차의 차이가 3.69로, 정확도가 SMSPP 대비 상대적으로 약 45% 낮았다.

반면, SMMP (w. LID)의 경우, SMSPP에 비해 오차의 차이가 0.33로, 정확도가 SMSPP 대비 상대적으로 약 4% 낮아 두 모델의 정확도는 거의 비슷하였다. 즉, 도로구간의 구분자로서 도로구간의 링크 아이디 (LID) 를 이용하면 SMMP가 SMSPP와 비슷한 성능을 낼 수 있음을 확인하였다.

3.3 실험 2. 도로구간 별 예측 정확도 비교

실험 2에서는 링크 아이디를 이용한 SMMP와 SMSPP의 10개 도로구간 각각에 대한 정확도를 확인하였다.



(그림 3) SMMP와 SMSPP의 도로구간 별 예측 정확도.

결과는 그림 3과 같다. 그림 3의 x축 항목들은 링크 아이디를 의미하고 y축은 MAPE이다. 정확도는 모든 도로구간에서 큰 차이가 없다. SMSP에 비해 링크 아이디를 이용한 SMMP의 오차가 최대 2.4 높고 1.2 낮다.

이를 통해, 전체 정확도와 마찬가지로 링크 별 정확도 또한 특정 도로구간에서 크게 하락하지 않고 SMSP와 비슷하게 보존된다는 사실을 검증하였다.

3.4 실험 3. 예측 수행 시간 비교

실험 3에서는 두 모델이 예측하는 데에 소요된 시간을 비교한다. 링크 아이디를 이용하는 SMMP가 SMSP에 비해 적은 수의 모델을 사용하기 때문에 예측을 수행하는 시간이 적을 것을 기대하였다.

측정 시간은 첫 번째 링크부터 마지막 링크까지 주어진 횟수만큼 예측하는 데에 소요된 시간이다. 즉, 차례로 첫번째 링크 속도 x회 예측, 두번째 링크 속도 x회 예측, ..., 10번째 링크 속도 x회 예측을 모두 마친 시간을 측정하였다.

<표 1> 예측 수행 시간 비교.

	1회	100회	10000회
SMMP	3.69초	6.83초	305.58초
SMSP	20.71초	24.33초	322.30초

표 1의 1회, 100회, 10000회는 도로구간 별 예측 횟수를 의미한다.

시간은 SMMP가 모두 약 17초정도 빠르다. SMSP는 다른 도로의 속도를 예측할 때마다 해당 도로의 데이터를 학습한 모델을 메모리에 로드 하여야 한다. 반면 SMMP는 한번에 한 개의 모델을 메모리에 로드하고 이를 계속 사용한다. 다시 말해, 17초는 모델이 메모리에 로드되는 시간에 의해 차이가 발생하는 것으로 나타났다.

다수의 모델을 병렬적으로 메모리에 로드 하더라도 실제로는 10개보다 훨씬 많은 도로가 존재하기 때문에 물리적 컴퓨팅 자원의 한계가 있다. 대한민국의 경우, 고속도로 상의 도로구간의 수는 3,000,000개 이상이다. 따라서, 링크 아이디를 이용한 SMMP는 정확도를 보존하면서 메모리에 모델을 로드하는 비용을 크게 줄일 수 있기 때문에 SMSP보다 실용적이라 할 수 있다.

4. 결론

본 논문에서는 딥 러닝을 이용한 다중 도로구간

속도 예측을 수행하는 단일 모델을 제안하고 그 성능을 실제 고속도로의 속도 데이터를 통해 검증하였다. 그 결과 링크 아이디를 입력 데이터로 활용하여 단일 도로구간 속도 예측 모델과 비슷한 정확도를 보였고, 예측 수행시간을 크게 단축시켰다.

Acknowledgements

이 논문은 정부(과학기술정보통신부)의 재원으로 (1) 한국연구재단-차세대정보·컴퓨팅기술개발사업 (No. NRF-2017M3C4A7069440), (2) 한국연구재단 (No.2018R1A5A7059549), 그리고 (3) 현대자동차 그룹의 재원으로 지원받아 수행된 연구임.

참고문헌

- [1] Eleni I. Vlahogianni, Matthew G. Karlaftis and John C. Golias “Short-term traffic forecasting: Where we are and where we’re going,” *Transportation Research Part C: Emerging Technologies*, Vol. 43, No. 1, pp. 3–19, 2014
- [2] N. L. Nihan and K. O. Holmesland, “Use of the Box and Jenkins time series technique in traffic forecasting,” *Transportation*, Vol. 9, No. 2, pp. 125–143, 1980.
- [3] X. Fei, C. Lu, and K. Liu, “A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction,” *Transportation Research Part C: Emerging Technologies*, Vol. 19, No. 6, pp. 1306–1318, 2011.
- [4] C.-H. Wu, J.-M. Ho, and D.-T. Lee “Travel time prediction with support vector regression,” *IEEE transactions on Intelligent Transportation Systems*, Vol. 5, No. 4, pp. 1438–1442, 2004.
- [5] Yisheng Lv et al., “Traffic Flow Prediction With Big Data: A Deep Learning Approach,” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 2, pp. 865–873, 2015.
- [6] Yuhan Jia, Jianping Wu, Yiman Du et al. “Rainfall-integrated traffic speed prediction using deep learning method,” *IET Intelligent Transport Systems*, Vol. 11, No. 9, pp. 531–536
- [7] Jaewon Choe et al., “On Applying Environmental Factors to Accurate Traffic Speed Prediction,” *KDBC, Chuncheon, Korea*, 2019, pp. 41–43.
- [8] Xing-Yu Chen, Hsing-Kuo Pao, Yuh-Jye Lee, “Efficient traffic speed forecasting based on massive heterogenous historical data,” *IEEE International Conference on Big Data*, Washington, DC, USA, 2014, pp. 10–17
- [9] Ibai Lana, Javier Del Ser, Manuel Velez, Eleni I. Vlahogianni, “Road Traffic Forecasting: Recent Advances and New Challenges,” *IEEE Intelligent Transportation Systems Magazine*, Vol. 10, No. 2, 2018.
- [10] Xiaolei Ma, Zhimin Tao, Yin Hai Wang, Haiyang Yu, Yunpeng Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data,” *Transportation Research Part C: Emerging Technologies*, Vol. 54, pp. 187–197, 2015.