

정확한 교통 속도 예측을 위한 환경요소의 적용*

최재원⁰¹, 이시영¹, 김남혁², 송준호¹, 박성환³, 김상욱^{†1}

¹한양대학교 컴퓨터소프트웨어학과

²한양대학교 자동차전자제어공학과

³현대자동차

{dnjsdlghdi, yurisector, huk320, hahadori, wook}@hanyang.ac.kr,

sunghwan.park@hyundai.com

On Applying Environmental Factors to Accurate Traffic Speed Prediction

Jaewon Choe⁰¹, Lee Siyoung¹, Namhyuk Kim², Junho Song¹, Sunghwan Park³, Sang-wook Kim^{†1}

¹Department of Computer Science, Hanyang University

²Department of Automotive Electronics and Control Engineering, Hanyang University

³Hyundai Motor Company

요 약

최근 교통 속도를 예측하고자 하는 연구가 활발히 이루어지고 있다. 하지만, 기존 대부분의 연구는 날씨와 시간 등의 환경요소를 고려하지 않고 과거 교통 속도를 활용하여 문제를 해결하고자 하였다. 본 논문에서는 기존 과거 속도만 활용하는 예측 모델에 환경요소를 추가적으로 적용하여 속도 예측 정확도에 미치는 영향력을 검증하였다. 예측 모델은 최근 널리 활용되고 있는 서포트 벡터 머신 (support vector machine)과 심층학습 (deep learning)을 선정하였고, 실제 서울시 교통 속도 데이터를 활용하여 검증하였다. 그 결과, 환경요소 데이터에 의해 예측 정확도가 향상됨을 확인하였다.

1. 서 론

특정 도로를 주행하는 차량들의 평균 속도인 교통 속도는 최적화 경로 제공, 교통 신호 체계 최적화, 실시간 돌발상황 대응 등을 종합적으로 관리하는 지능형 교통 체계 (ITS, Intelligent Transport Systems)에서 중요한 정보로 활용된다. 따라서, 최근 교통 속도를 예측하는 연구가 활발히 이뤄지고 있다 [1, 2].

기존 교통 속도 예측 관련 연구는 특정 기간의 과거 교통 속도를 기반으로 미래의 교통 속도를 예측하는 연구가 대부분이었다 [3-9]. 그러나, 교통 속도는 과거 교통 속도뿐만 아니라 교통 환경의 영향도 받기 때문에, 이를 고려하여 예측할 필요가 있다. 예를 들어, 교통 속도는 맑은 날보다 비 오는 날에 대체로 낮아지는 경향이 있고, 평일 교통 속도는 출퇴근 시간에 주로 낮아지지만 공휴일에는 같은 시간에 원활한 교통흐름을 보이는 경향이 있다.

따라서, 본 연구에서는 교통 환경 요인이 교통 속도 예측에 미치는 영향을 검증한다. 이를 위해, 교통 속도에 영향을 미치는 요인을 날씨, 이벤트, 시간 세 가지 항목의 환경요소로 정의한다. 여기서 날씨 요인은 기온과 강수량, 이벤트는 도로 위 사고와 공사 현황, 시간은 요일 특성, 월 항목을 포함한다. 환경요소별 속도 예측에 미치는 영향력을 검증하기 위해, 그림 1과 같이 교통 속도 예측 분야에서 널리 활용되는 서포트 벡터 머신 모델 (support vector machine)과 심층학습 모델 (deep learning), 두 가지 모델을 선정하여 정확도를 비교하였다 [10, 11].

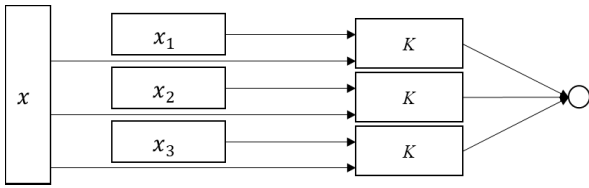
데이터 수집을 위해, 서울시 고속 도로의 일부 구간을 선정하여 교통 속도 및 환경요소 데이터를 수집하였고, 과거 속도와 각 환경요소의 조합에 따른 속도 예측 정확도를 비교하였다. 그 결과, 과거 속도와 환경요소를 모두 적용하였을 때, 서포트 벡터 머신 모델은 약 4%, 심층학습 모델에서는 약 8%의 예측 정확도가 향상되었다. 두 모델에서 모두 예측 정확도가 향상된 것으로 볼 때, 환경요소가 교통 속도 예측 정확도 향상에 기여한다는 것을 확인하였다.

2. 문제 정의

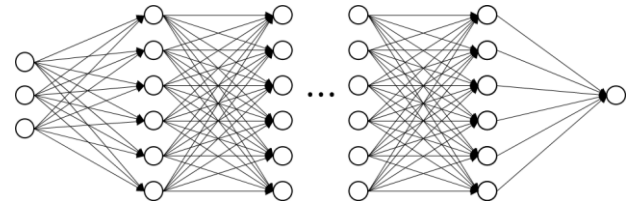
과거 속도 및 환경요소를 입력하여 한 시간 후의 속도를 예측하는 모델을 통해 환경요소 조합에 따른

* 이 논문은 정부 (과학기술정보통신부)의 한국연구재단-차세대정보·컴퓨팅기술개발사업 (No. NRF-2017M3C4A7069440), 과학기술정보통신부와 정보통신기획평가원의 SW 중심대학지원사업 (2016-0-00023) 및 현대자동차 그룹의 재원으로 지원을 받아 수행된 연구임.

† 교신저자



(a) 서포트 벡터 머신 (SVM, Support vector machine)



(b) 심층학습 (Deep learning)

그림 1. 예측 모델 구조

항목	값	설명
Temperature	0~1	기온 값을 0~1의 범위로 정규화
Rainfall	0~1	강수량 값을 0~1의 범위로 정규화
Accident	0, 1	사고 발생한 경우 1 표기
Construction	0, 1	공사 중인 경우 1 표기
Day	1~7	월요일부터 일요일을 순서대로 1~7 표기
Month	1~12	월을 그대로 표기
Workday	0, 1	공휴일 제외한 모든 날 1 표기
Holiday	0, 1	공휴일 1 표기
Post-holiday	0, 1	다음날이 공휴일인 경우 1 표기
Pre-holiday	0, 1	전날이 공휴일인 경우 1 표기

표 1. 환경요소 정의 및 구성

예측 정확도 향상을 가정하고 검증한다.

입력 데이터는 과거 속도 데이터와 날씨, 이벤트, 시간 데이터로 이루어진다. 과거 속도 데이터는 예측하고자 하는 도로의 과거 60분부터 현재까지 5분 단위로 관측한 값으로 구성된 벡터이다. 날씨 데이터는 과거 속도 데이터와 같은 기간에 관측된 기온 값과 강수량 값의 벡터이다.

이벤트 데이터는 과거 속도 데이터와 같은 기간 동안 사고나 공사 발생 여부를 표기한 one-hot 벡터이다. 또한, 예측하고자 하는 날의 다음날과 전날이 공휴일인지에 따라서도 교통 속도 양상이 달라질 수 있기 때문에, 이를 고려하기 위해 시간 데이터를 요일 특성과 월 벡터로 구성하였다. 환경요소 정의 및 구성은 표 1과 같이 정리하였다.

3. 실험

3.1. 실험 환경

실험 대상 도로는 경부 고속 도로 반포 IC 부근 하행선으로 선정하였고, 2018년 7월부터 10월까지의 교통 속도와 환경요소 데이터를 활용하였다.

5겹 교차 검증 (5-fold cross validation)으로 학습 데이터와 실험 데이터를 나누어 실험하였으며, 서포트 벡터 머신 모델의 커널 함수는 RBF (radial basis function)로 선정하였다. 심층학습 모델은 Fully-

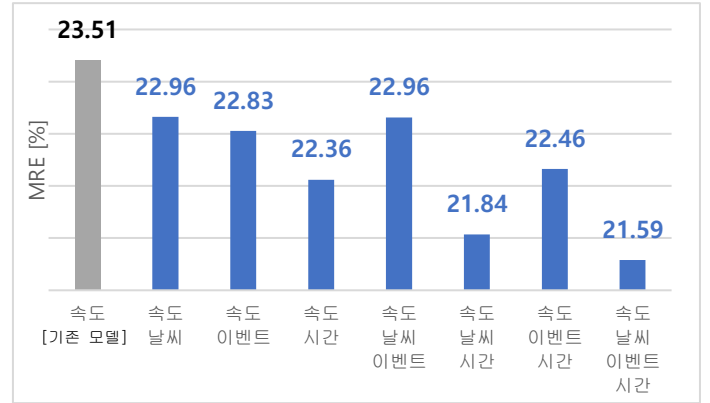
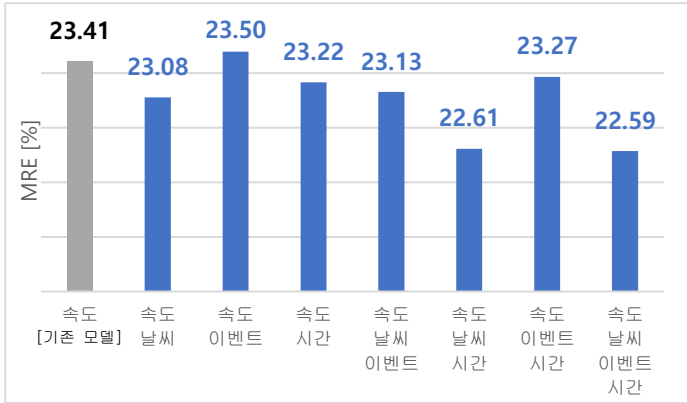
connected networks를 활용하였다. 은닉 노드의 활성화 함수는 Relu를 사용하였고, 최적화를 위해 Adam optimizer를 사용하였다.

3.2. 실험 결과

그림 2는 환경요소 적용에 따른 예측 상대오차 (mean relative error)이다. 두 모델 모두 환경요소 데이터가 추가적으로 적용됨에 따라 전반적으로 상대오차가 낮아지는 결과를 얻었고, 날씨, 이벤트, 시간의 항목을 모두 적용한 경우가 가장 낮은 상대오차를 보였다. 과거 속도만 고려하여 예측하였을 때, 서포트 벡터 머신 모델은 23.41%, 심층학습 모델은 23.51%의 상대오차를 보였다. 환경요소를 모두 적용하여 예측하였을 때는 각 22.59%, 21.59%로 감소하였다. 따라서, 예측 정확도는 약 3.5%, 8.2% 향상되었고, 심층학습 모델이 서포트 벡터 머신 모델보다 환경요소에 의해 예측 정확도가 더 큰 폭으로 향상되었다.

단일 환경요소를 추가하였을 때, 서포트 벡터 머신 모델은 날씨 데이터가, 심층학습 모델에서는 시간 데이터가 가장 큰 정확도 향상을 보였다. 한 시간 뒤 교통 속도는 기온, 강수량과 같은 날씨 데이터와 요일 특성과 같은 시간 데이터의 영향력이 크다는 것을 알 수 있다. 예를 들어, 오전 8시의 교통 속도는 예측하고자 하는 날이 평일이라면 출퇴근 시간과 겹쳐 도로에 따라 정체 현상으로 낮은 값이 될 가능성이 높을 것이고, 공휴일이라면 상대적으로 높은 값이 될 가능성이 높을 것이다.

반면, 이벤트 데이터의 경우, 서포트 벡터 머신 모델에서 영향력이 거의 없었고, 심층학습 모델에서는 일부 미미한 향상을 보였다. 서포트 벡터 머신 모델에서는 이벤트 데이터만 추가했을 때, 정확도가 오히려 감소하였고, 심층학습 모델에서도 날씨 데이터를 추가한 모델에 이벤트 데이터를 추가하여도 정확도 향상이 거의 없었다. 이는 데이터 수집기간의 사고와 공사 정보는 다른 항목에 비해 드물게 발생하여 대부분은 이벤트가 발생하지 않은 0으로 구성된다. 따라서, 이벤트 데이터로써 다수의 0만 추가로 입력되는 경우가 대부분이기 때문에 데이터의 차원만



(a) 서포트 벡터 머신 (SVM, Support vector machine)

(b) 심층학습 (Deep learning)

그림 2. 환경요소 적용에 따른 상대오차

증가시키고 최종적으로 이벤트 데이터를 추가하지 않았을 때와 성능 차이가 거의 없거나 데이터 차원 증가로 인한 정확도 감소가 발생한 것으로 판단된다.

하지만 환경요소 데이터를 모두 추가했을 때의 예측 정확도가 가장 큰 폭으로 향상되었기 때문에, 이벤트 데이터는 다른 환경요소 데이터의 조합에 따라 예측 정확도에 미치는 영향력이 달라지는 것을 알 수 있다.

4. 결 론

환경요소를 적용한 두 가지 교통 속도 예측 모델을 활용하여 환경요소 항목별 영향력을 검증하였다. 이를 위해, 각 데이터 조합에 따른 예측 정확도를 비교한 결과, 서포트 벡터 머신 모델은 단일 항목으로 날씨 데이터가 예측 정확도에 미치는 영향력이 가장 컸고, 심층학습 모델은 시간 데이터가 영향력이 가장 컸다. 반면, 두 모델 모두 이벤트 데이터의 영향력은 적었고, 환경요소를 모두 적용한 경우에 예측 정확도가 가장 큰 폭으로 향상되었다. 또한, 예측 모델로 널리 활용되는 서포트 벡터 머신과 심층학습을 선정하여 실험하였고, 이는 다양한 교통 속도 예측 모델에서 환경요소 데이터를 추가로 제공함으로써 예측 정확도를 향상시킬 가능성을 보였다. 따라서, 교통 속도 예측 문제에서 과거 속도와 더불어 환경요소를 입력 데이터로 구성하면 예측 정확도 향상을 기대할 수 있다.

5. 참조문헌

[1] E. Vlahogianni, M. Karlaftis, and J. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, Vol. 43, pp. 3-19, 2014.

[2] I. Laña et al., "Road Traffic Forecasting: Recent Advances and New Challenges," *IEEE Intelligent Transportation Systems Magazine*, pp. 93-109, 2018.

[3] Y. Rajabzadeh, A. Rezaie, and H. Amindavar,

"Short-term traffic flow prediction using time-varying Vasicek model," *Transportation Research Part C: Emerging Technologies*, Vol. 74, pp. 168-181, 2017.

[4] 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.

[5] Y. Zhang and Y. Liu, "Traffic forecasting using least squares support vector machines," *Transportmetrica*, Vol. 5, No. 3, pp. 193-213, 2009.

[6] C. Wu, J. Ho, and D. Lee, "Travel-time prediction with support vector regression," *IEEE transactions on intelligent transportation systems Syst.*, Vol. 5, No. 4, pp. 276-281, 2004.

[7] Y. 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, 2014.

[8] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, Vol. 79, pp. 1-17, 2017.

[9] Y. Wu et al., "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies*, Vol. 90, pp. 166-180, 2018.

[10] C. Cortes and V. Vapnik. "Support-vector networks." *Machine learning*, Vol. 20, pp. 273-297, 1995.

[11] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, Vol. 521, pp. 436-444, 2015.