**References**

1. Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic Flow Prediction With Big Data: A Deep Learning Approach," \*IEEE Transactions on Intelligent Transportation Systems\*, vol. 16, no. 2, pp. 865-873, Apr. 2015.

1. H. Zheng, H. Yao, and X. Zhu, "Short-Term Traffic Flow Prediction: A Deep Learning Approach," \*IEEE Transactions on Intelligent Transportation Systems\*, vol. 20, no. 6, pp. 2129-2139, June 2019.

4. X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. 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, May 2015.

5. L. Sun, D. W. C. Ho, J. Xu, and C. K. Wu, "Optimal scheduling of dynamic demand responsive transport services with vehicle overlap," \*IEEE Transactions on Intelligent Transportation Systems\*, vol. 15, no. 5, pp. 1928-1939, Oct. 2014.

6. B. C. Williams and L. A. Hoel, "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results," \*Journal of Transportation Engineering\*, vol. 129, no. 6, pp. 664-672, Nov. 2003.

7. L. An and Y. Peng, "Traffic flow forecasting based on random subspace and neural networks," \*IEEE Transactions on Intelligent Transportation Systems\*, vol. 16, no. 4, pp. 2063-2068, Aug. 2015.

8. G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," \*Science\*, vol. 313, no. 5786, pp. 504-507, July 2006.

9. V. Mnih and G. E. Hinton, "A Scalable Hierarchical Distributed Language Model," in \*NIPS\*, vol. 21, 2009, pp. 1081-1088.

10. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," \*Neural Computation\*, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

11. A. Graves, A.-r. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," in \*ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings\*, 2013, pp. 6645-6649.

12. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in \*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)\*, Honolulu, HI, 2017, pp. 1800-1807.

13. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in \*2014 IEEE Conference on Computer Vision and Pattern Recognition\*, Columbus, OH, 2014, pp. 580-587.

14. S. J. Pan and Q. Yang, "A Survey on Transfer Learning," \*IEEE Transactions on Knowledge and Data Engineering\*, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.

15. X. Li, G. Pan, Z. Wu, G. Qi, S. Li, D. Zhang, and W. Zhang, "Prediction of urban human mobility using large-scale taxi traces and its applications," \*Frontiers of Computer Science\*, vol. 6, no. 1, pp. 111-121, Feb. 2012.