# Abstract

In the traffic prediction field, deep learning models can be highly accurate, but they require large amounts of data to work correctly. This drawback called the “cold-start” problem, can prevent cities from starting their intelligent networks. To handle the situation, new models based on the concept of transfer learning were proposed. These models can learn complex spatio-temporal patterns from data-rich cities and transfer them to data-scarce counterparts. In this work, we aim to further develop the traffic prediction field by adapting elements of state- of-the-art models into a single model, taking advantage of external data, and using multiple cities as sources of knowledge.

# Acknowledgments

Acknowledgments go here

# Contents

1. [Introduction](#_bookmark1) 7
   1. [Motivation](#_bookmark2) 7
   2. [Research Questions](#_bookmark3) 7
   3. [Contribution](#_bookmark4) 8
   4. [Outline](#_bookmark5) 8
2. [Literature Review](#_bookmark6) 9
   1. [Traffic Forecasting](#_bookmark7) 9
   2. [Transfer Learning](#_bookmark8) 10
3. [Methodology](#_bookmark9) 11
4. [Results](#_bookmark10) 12
5. [Discussion](#_bookmark11) 13
6. [Conclusion](#_bookmark12) 14
   1. [Section](#_bookmark13) 14
7. [Tables](#_bookmark15) 15

List of Figures

[Figure 1 Title, Author](#_bookmark14) 14

# List of Tables

[Table 1 Caption](#_bookmark16) 15

[Table 2 Caption](#_bookmark17) 15

[Table 3 Caption](#_bookmark18) 15

# Introduction

## Motivation

According to [[1],](#_bookmark19) a smart city can be defined as a well-coordinated system that integrates advanced technological infrastructure, relying on sophisticated data processing. The primary objectives of such integration are to enhance city governance efficiency, improve citizen sat- isfaction, foster business prosperity, and promote environmental sustainability. Within a smart city, the management of various individual systems that constitute the urban environment is not solely reliant on the data collected within the city; it also relies on different adaptable models that can learn and evolve to suit the specific needs and characteristics of the city. In this context, the development of traffic prediction models emerges as a pivotal component for establishing the foundational framework of smart city management.

Recent advances in deep learning have led to significant advancements in prediction tasks related to traffic, such as crowd flow [[2,](#_bookmark20) [3],](#_bookmark21) traffic flow [[4,](#_bookmark22) [5],](#_bookmark23) public transit flow [[6,](#_bookmark24) [7],](#_bookmark25) travel demands [[8],](#_bookmark26) and traffic speeds [[9].](#_bookmark27) While formidable in their predictive power, these models come with a substantial data appetite. This data requirement poses a challenge for initiating new intelligent networks because meaningful inferences remain elusive despite the consid- erable investment needed to establish the sensor network without access to substantial data history. This difficulty is known in the field as the “cold-start” problem.

To address the aforementioned challenge, novel techniques rooted in transfer learning [[10]](#_bookmark28) have been introduced. These approaches enable training predictive traffic models for cities constrained by limited data by taking advantage of patterns observed in cities with abun- dant data resources. The fundamental concept behind these models involves the application of Multi-task Learning, a type of Inductive Transfer Learning, which entails initializing the network in the source city and implementing fine-tuning to adapt the network to the unique characteristics of the target city.

## Research Questions

The main objective of this work is to analyze and explore state-of-the-art models for traffic pre- diction with the intent of enhancing their accuracy. As a matter of organization, the following secondary objectives were drawn:

* Analyze current state-of-the-art models and identify cells and architectures that could be used when building a novel model;
* Based on the results of the first objective, propose a novel model to be built, which should

follow these requirements:

 capable of learning from multiple cities at the same time;

 capable of taking external data (such as weather data, [POI (Point of Interest)](#_bookmark0) locations, holidays’ calendar) as a supplement; and

 with a pre-processing step that includes a data augmentation cell.

* Analyze how impactful each feature derived from a requirement is to the model.

Concurrently, the following research questions were raised:

* 1. Is it possible to encompass more than two cities as sources in a transfer learning pro- cess?
  2. What’s the impact of the number of sources on the model’s accuracy?
  3. Is there a limit on the number of sources?
  4. Is data augmentation possible in the traffic prediction field?

## Contribution

To be done!

## Outline

This work is organized as follows: Chapter [2](#_bookmark6) introduces the literature review that was per- formed in order to further understand the problem and provides a comprehensive explanation of commonplace concepts of the field. Chapter [3](#_bookmark9) proposes a methodological framework for the entire work, including data acquisition, pre-processing, model building, and testing setup. Chapter [4](#_bookmark10) analyzes the results of the proposed tests and comparisons to proposed baselines. Chapter [5](#_bookmark11) uses the results to answer and discuss the research questions raised on Chapter

[1.](#_bookmark1) Chapter [6](#_bookmark12) concludes the thesis and discusses the main directions for future research in the field.

# Literature Review

This chapter presents a literature review that we conducted to further the traffic forecasting field and its state-of-the-art.

## Traffic Forecasting

Traffic Forecasting is a long-lasting field of study in Traffic Engineering, conceived in the 1950s [[11,](#_bookmark29) [12].](#_bookmark30) Initially centered on traffic simulation, the area observed a significant upward trend in recent years. By its very nature, a traffic network constitutes a vast and complex system where events occurring at various junctures within the road grid can exert profound influence over the entire traffic flow. These inherent complexities render it an ideal subject for examination by cutting-edge machine-learning algorithms, such as [LSTM (Long Short-Term](#_bookmark0) [Memory),](#_bookmark0) [GRU (Gated Recurrent Unit),](#_bookmark0) and [CNN (Convolutional Neural Network).](#_bookmark0)

Traffic network problems inherently comprise two domains: the temporal and spatial domains. In the early stages of deep learning model development for the field of traffic, a natural division emerged, allocating separate components to address each of these domains. For example, [CNN](#_bookmark0) has conventionally been harnessed primarily for spatial feature extraction, which in- volves identifying the physical locations or arrangements of elements within a given dataset. Conversely, [RNN (Recurrent Neural Network)](#_bookmark0) has specialized in temporal feature extraction, focusing on discerning patterns that evolve or change over time.

The pioneering work of [[13]](#_bookmark31) marked one of the earliest instances of employing [CNNs](#_bookmark0) for traf- fic prediction tasks. Subsequently, this methodology proliferated across various model frame- works and underwent integration with other architectural paradigms. An illustrative example of this evolutionary trajectory emerges from the study conducted by [[14],](#_bookmark32) which employed a multiple [GCN (Graph Convolutional Network)](#_bookmark0) framework. In this network type, the input to the convolutional layers consists of the city’s graph representation.

Similarly, [[15]](#_bookmark33) were among the early proponents of employing [LSTM](#_bookmark0) cells for short-term traffic prediction. This approach gained substantial traction, leading to the integration of cells into various models, as exemplified by the [ConvLSTM (Convolutional Long Short-Term Memory)](#_bookmark0) module introduced by [[16].](#_bookmark34) In this architecture, spatial features were extracted at each time frame and subsequently fed into an [LSTM](#_bookmark0) chain. This innovative approach allowed for the extraction of spatial and temporal features concurrently.

## Transfer Learning

Transfer learning techniques [[10]](#_bookmark28) were introduced as valuable tools for addressing problems where existing knowledge or expertise from one domain could be employed to enhance learn- ing and performance in another domain. These techniques prove incredibly beneficial when the latter domain suffers from a shortage of data.

# Methodology

Intelligent transportation systems (ITS) are

# Results

Intelligent transportation systems (ITS) are

# Discussion

Intelligent transportation systems (ITS) are

# Conclusion

Intelligent transportation systems (ITS) are

## Section

* + 1. **Subsection**

Example of a figure, see below.



**Figure 1** Title, Author

Example of glossary use: [ITS (Intelligent Transportation System).](#_bookmark0) Example use of citation [[17].](#_bookmark35)

# Tables

## Table style 1

### Spalte 1 Spalte 2

|  |  |
| --- | --- |
| Nummer 1 | Nummer 2 |
| Nummer 1 | Nummer 2 |
| Nummer 1 | Nummer 2 |

**Table 1** Caption

## Table style 2

|  |  |
| --- | --- |
| **Spalte 1** | **Spalte 2** |
| Nummer 1 | Nummer 2 |
| Nummer 1 | Nummer 2 |
| Nummer 1 | Nummer 2 |

**Table 2** Caption

## Table style 3

### Spalte 1 Spalte 2

Nummer 1 Nummer 2

Nummer 1 Nummer 2

Nummer 1 Nummer 2

**Table 3** Caption

# Bibliography

1. C. Yin, Z. Xiong, H. Chen, J. Wang, D. Cooper, and B. David, “A literature survey on smart cities.,” *Sci. China Inf. Sci.*, vol. 58, no. 10, pp. 1–18, 2015.
2. J. Zhang, Y. Zheng, D. Qi, R. Li, X. Yi, and T. Li, “Predicting citywide crowd flows using deep spatio-temporal residual networks,” *Artificial Intelligence*, vol. 259, pp. 147–166, 2018.
3. W. Jin, Y. Lin, Z. Wu, and H. Wan, “Spatio-temporal recurrent convolutional networks for citywide short-term crowd flows prediction,” in *Proceedings of the 2nd International* *Conference on Compute and Data Analysis*, pp. 28–35, 2018.
4. 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.

1. Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, “A hybrid deep learning based traffic flow prediction method and its understanding,” *Transportation Research Part C: Emerging* *Technologies*, vol. 90, pp. 166–180, 2018.
2. Y. Liu, Z. Liu, and R. Jia, “Deeppf: A deep learning based architecture for metro passen- ger flow prediction,” *Transportation Research Part C: Emerging Technologies*, vol. 101,

pp. 18–34, 2019.

1. D. Chai, L. Wang, and Q. Yang, “Bike flow prediction with multi-graph convolutional net- works,” in *Proceedings of the 26th ACM SIGSPATIAL international conference on ad-* *vances in geographic information systems*, pp. 397–400, 2018.
2. X. Geng, Y. Li, L. Wang, L. Zhang, Q. Yang, J. Ye, and Y. Liu, “Spatiotemporal multi- graph convolution network for ride-hailing demand forecasting,” in *Proceedings of the* *AAAI conference on artificial intelligence*, vol. 33, pp. 3656–3663, 2019.
3. B. Yu, H. Yin, and Z. Zhu, “Spatio-temporal graph convolutional networks: A deep learn- ing framework for traffic forecasting,” *arXiv preprint arXiv:1709.04875*, 2017.
4. 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, 2009.
5. M. J. Beckmann, C. B. McGuire, and C. B. Winsten, *Studies in the Economics of Trans- portation*. Santa Monica, CA: RAND Corporation, 1955.
6. H. W. Bevis, “A model for predicting urban travel patterns,” *Journal of the American* *Institute of Planners*, vol. 25, pp. 87–89, May 1959.
7. X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, “Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction,” *Sensors*, vol. 17, p. 818, Apr. 2017.
8. J. Lin, Y. Chen, H. Zheng, M. Ding, P. Cheng, and L. Hanzo, “A data-driven base station sleeping strategy based on traffic prediction,” *IEEE Transactions on Network Science* *and Engineering*, 2021.
9. Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, “LSTM network: a deep learn- ing approach for short-term traffic forecast,” *IET Intelligent Transport Systems*, vol. 11,

pp. 68–75, Feb. 2017.

1. H. Yao, F. Wu, J. Ke, X. Tang, Y. Jia, S. Lu, P. Gong, J. Ye, and Z. Li, “Deep multi- view spatial-temporal network for taxi demand prediction,” in *Proceedings of the AAAI* *conference on artificial intelligence*, vol. 32, 2018.
2. L. Wang, D. Chai, X. Liu, L. Chen, and K. Chen, “Exploring the generalizability of spatio- temporal traffic prediction: Meta-modeling and an analytic framework,” *IEEE Transac- tions on Knowledge and Data Engineering*, vol. 35, pp. 3870–3884, Apr. 2023.