Revolutionizing Last-Mile Delivery: Integrating Social Media and Deep Learning for Optimized Traffic Prediction in E-Commerce

Valeria Laynes Fiascunari, Luis Rabelo
Department of Industrial Engineering and Management Systems
University of Central Florida
Orlando, Florida, USA
va584101@ucf.edu

Edgar Gutiérrez-Franco
Center for Transportation and Logistics
Massachusetts Institute of Technology
Cambridge, MA, USA

Abstract

Effective traffic prediction is crucial due to a surge in deliveries by commerce and urbanization. This has led to a notable rise in traffic within megacities, causing route delays to the final destinations and countless vehicle accidents. E-commerce has been in a constant boom, as buying something online and having it delivered to the front door is easier than going to the store. As more people engage in this activity, e-commerce platforms' challenges are more complicated and need to be addressed faster. However, these challenges escape the delivery company's scope when external factors influence the objective of optimized deliveries, for example, traffic issues or bad weather during the last mile, issues that are only exacerbated where traffic sensors are not widely used (i.e., underdeveloped countries). The main contributions of this research are to (1) provide a contextual foundation of current frameworks used for traffic prediction, (2) use social media and multi-modal traffic-related data (weather, points of interest, calendar of events) by leveraging social network analysis to improve the accuracy of traffic prediction, and (3) to show a methodology that can be used for partially observed traffic. The proposed methodology includes deep learning tools like Long-Short Term Memory Networks, attention mechanisms, Graph Convolutional Networks, and social media tools like sentiment analysis.

Keywords
Social networks, deep learning, underdeveloped countries, traffic prediction.

1. Introduction

Urbanization is a direct consequence of population growth. Cities with more than 1 million inhabitants will rise by almost 30% from the year 2018 (548 habitants) to 2030 (706 habitants), and the number of megacities with more than 10 million inhabitants is “projected to rise from 33 in 2018 to 43 in 2030” (UN, 2018). These megacities are located all over the world, from New York in North America and London in the United Kingdom to Lima in Peru, Bogota in Colombia, and Mexico City in Mexico. However, these are generally located in the southern hemisphere. Population growth would not have a concerning impact on everyday life if it increased inhabitants in rural areas, but that is not the case. Population density in urban areas has always been increasing, but since 2010, urban areas have a higher
Given the amount of increasing population density and urbanization, vehicle ownership is also impacted consequently. The results of studies done on the relationship between urbanization and vehicle ownership have two outcomes. On the one hand, vehicle ownership decreases in developed countries when the urban areas are further developed (for example, construction of points of interest or intersections) (Sabouri, Tian, Ewing, Park, & Greene, 2021), meaning that people do not see the need of owning a vehicle due to the proximity to jobs or schools and advanced public transportation. On the other hand, vehicle ownership increases on underdeveloped countries when roads, parking lots, and vehicle-friendly structures are built, increasing vehicle density within the city (Anirudh, Mazumder, & Das, 2022), which creates a slippery slope fallacy where the more the city is developed, the higher the vehicle density will be.

Smart city planning has included population growth, urbanization, and vehicle ownership to build infrastructure like tunnels, road upkeeping, and reliability checks. Other plans include sustainable transportation, reduction of gas emissions, scrapped vehicle recycling, among other advanced projects to aid with the rapid density growth. One of the most researched topics within transportation systems is traffic prediction, which is used in planning. This includes adding IoT systems to increase data repositories (Bresciani, Ferraris, & Del Giudice, 2018) and bring added value to citizens while improving the economy by bringing innovative companies and their solutions to a city-wide integration. While this, in turn, brings fast-evolving research venues like smart city planning (Axelsson & Granath, 2018) or policies on urban innovation (Caraglì & Del Bo, 2019), underdeveloped countries are still trapped on a slippery slope, without infrastructure, resources, or the knowledge to fix a basic urban issue: tardiness due to traffic.

### 1.1 Traffic Management Challenges in Underdeveloped Countries

Latin America is the most urbanized region in the world due to rural exodus (Estupiñán et al., 2018). However, cities in this region are behind in Intelligent Transportation Systems despite having chaotic traffic. Lima in Peru, Bogota in Colombia, and Mexico City in Mexico are some examples, having the eighth, tenth, and thirteenth places, respectively, for the cities with the worst traffic congestion in the world. Furthermore, Lima and Bogota experienced an average time of traveling 10km of 27 min with 7 sec and 26 min with 20 sec in 2022, respectively (TomTom, 2023). On the other hand, cities in the United States have at most a travel time of 24 min with 30 sec (New York) (TomTom, 2023). The growth in urbanization and in congestion calls for an enhanced ITS. Despite the cities’ need for an update, there are still challenges that prevent Latin America from improving the system: lack of long-term mobility policies, resistance by transport operators to system integration, limited infrastructure, lack of financial resources, and limited use of bank accounts by the population (Toro, Krogt, & Vallejo, 2021). Unfortunately, some of these challenges are correlated. For example, if a city lacks financial resources, it cannot have a high-end infrastructure. Public transportation is widely available in the US, where cities like Los Angeles, New York, and Seattle have a multimodal system including buses, subways, light rails, etc. Improving public transportation to reduce the number of private vehicles on the road is an ongoing project in many Latin American countries, but budget limitations and lack of planning make this a multi-year project, where traffic is still a problem due to many short-distance routes like going to work, getting groceries, or taking the kids to school (Ramirez, 2023).

### 1.2 The Need for Effective Traffic Prediction

Industry grows fast for B2B and B2C e-commerce companies and management face a big challenge: bar keeps raising, and customers will always turn to the best option, usually the faster and most reliable service. Specifically, last-mile delivery has the urge to meet good service delivery and usually that leads to poor vehicle utilization and service duplication (Allen et al., 2018). They set a list of challenges that last-mile delivery service face like (1) demand patterns and its peak times, where processes are pushed nearly out of control, (2) times being shortened between the orders being placed and the delivery of the product by companies who optimize quickly, (3) sets of times too complex for the service to deliver quality, including time windows and restrictions, (4) delivery failures for residential addresses, (5) high return rate due to failures in attempts to manage stated challenges, and (6) lack of logistics, like distribution centers, budget, vehicles, drivers, etc. to fulfill the demand rate. Cities rely on traffic prediction for accident prediction, road repairs, environmental concerns, just as supply chain companies rely on it to optimize their delivery routes on last-mile deliveries. Many factors can be determined by either, cities or companies, that impact on traffic. However, external factors, such as weather, can also be implemented in routing optimization to avoid traffic congestion, which can be determined by many intelligent transportation systems like lane change behavior analysis, passenger flow analysis, vehicle combination effect on traffic flow, among others (Balas, Jain, & Zhao, 2017). A
report outlines many benefits that Intelligent Transportation Systems bring, among them are predicting traffic, which in fact, can make commercial vehicles more efficient, decrease traffic congestion in metropolitan areas and improve safety issues (Intelligent Transportation Systems: real world benefits, 1998).

1.3 Objectives
The need for effective traffic prediction extends beyond urban planning to impact customer satisfaction in the e-commerce industry. The accelerated growth of online shopping, driven by the demand for fast and reliable services, emphasizes the importance of accurate delivery predictions. For logistics service providers (LSPs), meeting customer expectations for reasonably priced, fast, and reliable services is crucial for customer loyalty. However, challenges in last-mile delivery, such as demand patterns, shortened delivery times, and high return rates, pose obstacles to achieving optimal service levels. Traffic prediction becomes a vital component for e-commerce companies and cities alike to enhance last-mile delivery efficiency, reduce congestion, and meet customer expectations. Despite various factors influencing traffic, challenges persist in accurately predicting traffic in residential areas due to the limited deployment of traffic sensors, especially in underdeveloped regions.

The following research questions were derived before conducting the systematic literature review and taking into consideration the challenges stated. These questions are the basis of the research and are included as eligibility criteria for deletion and inclusion of papers in the study and further analysis.

a) What are the current methods developed for traffic prediction using deep learning tools and external factors such as weather, social media, and points of interest?
b) How can multimodal data, including information from diverse sources such as traffic sensors, GPS devices, and points of interest, be effectively integrated to enhance the accuracy of traffic prediction models in underdeveloped countries, where limited infrastructure and resources present unique challenges?
c) Could this framework be an initial point to further implement a real-time route optimization, if needed, for highly variable optimal routes?

2. Literature Review
The objectives of the literature review are threefold: first, to provide a contextual foundation of current technologies used for spatio-temporal traffic prediction and their levels of accuracy, computational complexity, and data availability; second, to summarize the instances where social media has been used as vehicular traffic predictor and how the data has been handled; and third, explore the challenges that traffic prediction has in light of limited data availability in underdeveloped countries when using technologies from the first objective. Therefore, this literature review provides a background to identify research gaps where the methodology proposed can be applied.

2.1 Recurrent Neural Networks with Convolutional Neural Networks
Hybrid networks that use CNN and RNNs are mostly used with spatiotemporal data, where both spatial and temporal features are required to have a high accuracy. It merges the strengths of RNNs for temporal data and CNN for spatial data. In traffic prediction, this is one of the most used methods when the input data only includes one dataset for spatial data, like a network of traffic sensors deployed in a city, and one dataset for temporal data, like readings over time of each of these sensors. These types of models will usually outperform CNN (Cheng, Lu, Zhou, Zhang, & Zhang, 2022; Zang, Ling, Wei, Tang, & Cheng, 2019) and LSTM when used on their own (W. Lu, Rui, & Ran, 2020). Since it exploits each NNs characteristics, it can more accurately predict new sensors based on neighboring ones (Fouladgar, Parchami, Elmasri, Ghaderi, & Ieee, 2017) and have more customization in model architecture (Han, Chen, & Sun, 2019).

Weather implementation in road network traffic prediction holds the most advanced models for CNN-RNN architectures, which usually hold multidimensional feature space (Jingyuan Wang, Cao, Du, & Li, 2019; Z. Wang, Ding, & Wang, 2021). However, since the convolutions still hold a high-demand computation, some methodologies include residual networks to support more convolutional layers needed for high-dimensions (J. Zhao & Zhu, 2021). These studies could have improvements like being able to handle large road networks (J. Xu, Zhang, Jia, & Xing, 2019), using solely ResNet for spatial features (Elmi & Tan, 2020), or use it as part of the feature engineering phase to extract features (R. He, Liu, Xiao, Lu, & Zhang, 2022). Other methods of reducing complexity include incremental learning, which is special needed for a wider range of factors such as traffic, weather, and traffic accidents (Shao, Zhao, Yu, Zhu, & Fang, 2021).
CNN-RNN models have a good accuracy for low dimensional data, requiring help from other machine learning mechanisms every time another variable is added. As stated above, there are many factors that affect traffic patterns, some induce more uncertainty than others, and most of them are geographically placed. Since CNNs are limited to 3D, not all the influential factors are going to fit the model. The next section introduces a generalization of CNNs that was design for multidimensional graphs.

2.2 Neural Networks with Graph Methods

There are many types of neural networks that include graphs, like Graph Convolutional Network (GCN) and Graph Attention Network (GAN). Unlike CNN-RNN architectures, GNN can handle multi-modal data and can propagate the information across the nodes in a graph depending on the number of convolutions (Agafonov & Ieee, 2020). The input graph can be directed or undirected, can hold weights or other type of information on the edges (Y. Lv, Cheng, Lv, & Li, 2022) that could help improve graph representation. There is limited research for GNNs predicting traffic for one road, freeways or highways, this includes short-term and daily prediction at toll gates (Shi, Yuan, Wang, & Zhao, 2021), dividing data into components (seasonal, static, acyclic) (Y. S. Shen, Li, Xie, Li, & Xu, 2022), and pre-processing methods that improve accuracy like feature engineering (Mihaita, Papachatgis, & Rizoiu, 2020) and models for incomplete or missing data (Y. Liu, Wu, Wen, Xiao, & Chen, 2022). Highways are not always represented as a graph, while GNNs expect an explicit graph structure. Forcing a single-road system into a graph representation might not capture the spatial relationships optimally and, hence, not yielding the best results.

Inherently, GNNs are designed to process graph-like data, which make them useful for road-network traffic predictions. Some studies focus on pre-processing features for incomplete data (Zhong, Suo, Jia, Zhang, & Su, 2021), getting graph analysis metrics (Z. Q. Hu, Shao, & Sun, 2022; Q. Wu, Fu, & Nie, 2020), and correlating nodes that are not geographically close (M. Lv et al., 2021; Y. M. Zhao, Cao, Zhang, & Liu, 2021). These is a necessary step for GNNs since it must fit a graph form, which might require extensive engineering. Other studies increase accuracy by adapting a divide-and-conquer approach (Chattopadhyay & Tham, 2022), using techniques like clustering (Huang, Song, Zhang, & Yu, 2021; C. Zhang, Zhang, Yu, & Yu, 2020) and separating roadmap into sub-regions (Ren & Xie, 2019; N. Zhang, Guan, Cao, Wang, & Wu, 2019). These techniques let the model use knowledge transfer from one sub-graph to another or between graphs in the same cluster. The goal of this techniques is to reduce computational complexity by complementing the model’s learning process.

GNNs frameworks concentrating on the temporal feature for prediction accuracy is a common practice given that the data is more user-friendly. Including multiple components of periodicity with convolutions prioritizing temporal components (M. Liu et al., 2023; Yang, Wen, Yu, Zhang, & Ieee, 2020) and modeling asynchronous relation within a road network (Qi, Li, Chen, & Xue, 2021) are some examples. However, working with the spatial components has shown to yield better results. Construction of hypergraphs allows for more flexible relationships among nodes, some which may not be geographically related (F. X. Li et al., 2023; J. C. Wang et al., 2022; Y. Wang & Zhu, 2022) like a similarity graph with node embedding (Gou, Han, & Zhang, 2022) or multigraph convolutions for nodes further apart (J. Ye, Zheng, Zhao, Ye, & Xu, 2021). A two-component GCN considers global spatial relations between non-first order sensors, while a local component analyzes first-order ones (Feng, Huang, Shen, Shi, & Shi, 2022), this method avoids the smoothing problem of GCNs, where information from neighboring nodes has to be aggregated (“smoothed”), other methods are constant modifications of the graph data to fit an encoder-decoder architecture (Xie et al., 2020) and using GCN-LSTM structure where the k-order convolution process k neighboring nodes (Yan, Wang, Yu, Jin, & Zhang, 2021). Graph data has a high complexity, especially when working with hypergraphs, data tends to be over smoothed. With multi-graph approaches, generalization might arise as a problem, since not all networks have data for all available roads. Meaning, there might low accuracy predictions for roads with sparse data.

If parallelized correctly, graph networks can be a solution for traffic prediction tailored for last-mile problems. They show a high accuracy due to their ability of retaining complex relationships within the road network and scalability to bigger networks or a wider range of features. Their usage with other methods, like resnets and attention mechanisms, make them very customizable, depending on the prediction objective.

2.3 Neural Networks with Attention Methods

Attention mechanisms are a valuable tool for neural networks. They increase interpretability and performance by discerning the most important inputs to generate the best outputs. Highlighting the features that drive the model’s decision makes it easier to perform with a high accuracy. This attention mechanism can be embedded to identify
important temporal dependencies (Abdelraouf, Abdel-Aty, & Yuan, 2021; S. Zhang, Guo, Zhao, Zheng, & Chen, 2021) or spatial dependencies (Miao, Su, Fu, Chen, & Zang, 2022), or both (H. Hu, Lin, Hu, & Zhang, 2022; Zhihong Li, Xu, Gao, Wang, & Xu, 2022; G. J. Shen, Yu, Zhang, & Kong, 2021). Road-network traffic prediction with attention mechanisms has been a common practice for half a decade, from including them in simple RNN or CNN architectures (Do, H.L, B.Q, Z, & D, 2019; Mao, Huang, Lu, Chen, & Liu, 2022), to more complicated ones like Graph Attention Networks (Rajkumar & Jegatha Deborah) or hybrid RNN-GCN with attention weights (Buroni, Libichot, & Bontempi, 2021; Zhishuai Li et al., 2019). Using dynamic graphs could help adapt to evolving situations like traffic accidents or sudden surges in traffic due to construction (X. Luo, Zhu, Zhang, & Li, 2023). This has been shown to aid simultaneous short and long-term predictions (B. Li, Guo, Wang, Gandomi, & Chen, 2021; Lin, Ge, Li, Zeng, & Ieee, 2022) and GRU models by capturing long-term dependencies while dynamically adjusting weights to important time steps (X. Luo et al., 2023). A more in-depth work used dynamic spatiotemporal graphs, where temporal and spatial dependencies are represented by dynamic edges (Z. Fang, Pan, Chen, Du, & Gao, 2021). This research used a four-component model where each had a special attention to dynamic properties. Nonetheless, it lacks traffic-related variables like weather.

One of the main issues with attention mechanisms is that it’s prone to only focus on local features, ignoring the global dependencies. There are many ways to mitigate this problem, like feeding the model different compositions of the temporal feature (Duan et al., 2022; B. Sun, Zhao, Shi, & He, 2021; Xue, Zhao, & Han), visibility graphs that considers global context and lane interconnectedness (Zeng & Tang, 2022), or adding layers to the model that capture long-term and global trends (Chen, Han, Yin, & Cao, 2020; R. Luo, Song, Huang, Zhang, & Su, 2022; X. S. Wu, Fang, Liu, & Wu, 2021). The choice of mitigating strategy will depend on the objective of the prediction and study, like focusing on getting global dependencies for spatial components.

Even though Graph Attention Networks (Rajkumar & Jegatha Deborah) perform better than GCN (D. Li & Lasenby, 2021; Z. Pan et al., 2020) for a regular road-network, attention mechanisms are particularly useful when using multi-modal data. (Jin et al., 2021; Jichen Wang, Zhu, Sun, & Tian, 2021; X. Xu et al., 2020), where a hypergraph can be evaluated (H. Zhang, Liu, Tang, Xiong, & Ieee, 2020). When including weather, for example, research shows an improvement in prediction with using an attention mechanism in the temporal component (de Medrano & Aznarte, 2020; K. Wang et al., 2021) which enables an LSTM to capture high dependencies, but when including taxi, bike, and weather data, decomposing the temporal component with a multi-granularity approach shows a better performance (Ali, Zhu, & Zakarya, 2021; J. Liu, Qu, Chen, & Gong, 2022; X. Zhang, Huang, Xu, & Xia, 2020). This is because, as stated before, attention mechanisms have issues capturing the global context of the data, and including different decompositions help understand better the input. Another setback for attention mechanisms is higher complexity, but this can be mitigated as well by introducing parallelization of tasks, as shown in divide-and-conquer approaches where the road graph is divided in smaller sections for analysis (Long et al., 2023; Zheyi Pan et al., 2019; J. R. Sun, Peng, Jiang, Hong, & Sun, 2022; Y. Ye et al., 2023). Another solution is to fuse model components (S. Fang et al., 2022; L. Liu et al., 2021), which lets the attention mechanism work, for example, on the temporal component and then fuses it with the spatial component.

These types of neural networks have been shown to increase accuracy in traffic prediction by enabling models to capture complex spatial and temporal patterns, while including multi-modal data. Even though they have some downsides like lack of global context or complexity, these can be mitigated by using the correct architecture and data preprocessing. Attention mechanisms are used in most of traffic research articles analyzed in this systematic literature review, which makes it a state-of-the-art methodology for the latest six years.

### 2.4 Social Media in Traffic Prediction

Studies showing correlation between social media and vehicular traffic have been around for a decade, showing a decrease in prediction error by using tweets semantics as traffic indicators (J. He, Shen, Divakaruni, Wynter, & Lawrence, 2013), or using topics within tweets as one of many multi-source features (L. Lu, Jianxin, Feng, Jieping, & Jinpeng, 2018). These types of social media usage arise challenges like tweet ambiguity i.e., choosing “traffic” as one of the topics and not differentiating “communication traffic” from “vehicular traffic”, which elevate complexity. Using an advanced Natural Language Processing (NLP) model that transforms tweets into time-series records (Tsai et al., 2022) diminished this ambiguity, but also lacks the computational power to handle other factors. Limiting the number of topics to only traffic accidents (Liyoug, Vatekul, & Ieee, 2019; Yao & Qian, 2021) or nonrecurrent events (Essien, Petrounias, Sampaio, & Sampaio, 2021) also show an improvement in accuracy without complicating the model. Placing accidents or events within a road network is one solution to social media usage without text mining,
given the geocoding of tweets, but these types of matrices will have an even bigger sparsity problem than non-geocoded ones.

Geographically placed social media requests help predict traffic around hotspots (Liao et al., 2018; Rajkumar & Jegatha Deborah, 2020). These types of requests are user-specific and make it hard to predict people mass movement around an entire network. A study claimed that using tweets from user accounts related to transportation in the city helps the problem of complex data mining (Dayong, Longfei, Jianping, & Senzhang, 2018). By doing so, they find co-occurrence congestion patterns, but important information regarding social media users that react to this type of news is not evaluated, nor its influence on the social network activity. This type of study is done in (Laynes Fiascunari & Rabelo, 2022), where a deterministic information diffusion model is performed after influential accounts post tweets. The research showed that the independent cascade model does not have an impact of the study and leaves the possibility of other diffusion models, as well as multi-modal factors, being of higher relevancy.

Partially observed traffic prediction is a great improvement (X. Liu, Kong, Li, & Acm, 2016), especially for underdeveloped countries that are not data-rich and surely have the same last-mile needs from a smart-city. This aims to use location-based social media information to fill in roads without sensors, but its limitations on the amount of geocoded data still holds, which could be complemented by analyzing the social network and not only specific tweets’ semantics. An advancement has also been made regarding the amount of factors to potentially include in the model (Nguyen, Dao, & Zettsu, 2020), where four urban sensing sources – twitter, traffic, accidents, weather – were stored in a single 2d multi-layer image per time stamp. This model also has the ability to integrate perfectly more data sources but has not been tested for partially observed traffic or using social networks instead of sparse geolocated tweets.

Social media holds data that could improve the accuracy of traffic prediction, it has been shown to be correlated and an important factor. Even though its usage might not be a big accuracy improvement for smart cities, cities where traffic is alarming and lack sensor technology could complement many sources and get a better prediction to use in last-mile deliveries.

3. Methods
Three main deep learning tools were used to build multiple permutations of the models. LSTM, GCN, and an attention mechanism were stacked and paired in different ways to explore the computational complexity and their accuracy when predicting traffic.

3.1 Long-Short Term Memory Neural Network
Described in the Literature Review section, is a Recurrent Neural Network (RNN) type that avoids the backpropagated error to be blown up or decayed exponentially by inserting an LSTM cell, illustrated in Figure 1, instead of just a Tanh layer.

![LSTM cell in RNN](image)

The LSTM NN has three gates: (1) the input gate, which controls if the memory cell is updated, and (2) the forget gate, which decides if the memory must be set to 0. And (3) the output gate, which decides if the information on the
current state of the cell is made visible. The formulas for each of these gates can be found in Formulas (1), (2), and (3), respectively.

\[
i^*(t) = \sigma(W^i \cdot [h^*(t-1) \cdot x^t] + b^i) \quad (1)
\]

\[
f^*(t) = \sigma(W^f \cdot [h^*(t-1) \cdot x^t] + b^f) \quad (2)
\]

\[
o^*(t) = \sigma(W^o \cdot [h^*(t-1) \cdot x^t] + b^o) \quad (3)
\]

Where \( \sigma \) is the activation function, \( W \) represents the weights, \( h^*(t-1) \) is the input from the previous computation, \( x \) is the input from the first layer (blue circle), and \( b \) is the bias. This neural network can manage sequence-to-sequence (seq2seq) problems, using an encoder to translate the input sequence to a vector and a decoder to translate the vector to the output sequence. This scheme is mostly used in text translation, handwriting recognition, speech recognition, and in this case, traffic prediction.

3.1 Graph Convolutional Network

Unlike regular Convolutional Networks, GCNs can handle high-dimensional and non-Euclidean data. GCNs have vertices that send and receive messages to and from other vertices through edges, and at each vertex and layer, an aggregation function is performed. The number of convoluted layers determines the number of neighborhoods the message will travel through. The architecture of the network is represented in Figure 2.

![Graph convolutional network architecture](image)

**Figure 2.** Graph convolutional network architecture.

GCNs use forward propagation using an adjacency matrix to include graph information in the formula. Formula (4) belongs to forward propagation, where \( H^{(l+1)} \) represents the feature representation in layer \( l+1 \), \( \sigma \) is the activation function, \( W^l \) is the weights for layer \( l \), \( H^l \) is the feature vector for layer \( l \) and \( b^l \) represents the biases for layer \( l \). In contrast, Formula (5) is the adjusted formula for GCN.

\[
H^{(l+1)} = \sigma(W^l H^l + b^l) \quad (4)
\]

\[
H^{(l+1)} = \sigma(D^{-1/2} A D^{-1/2} W^l H^l + b^l) \quad (5)
\]

After deriving an adjacency matrix from the road network map and using a GCN, each road segment can have a set of variables that defines it, characterizing them with weather conditions, points of interest, and degree of change from traffic influencers. Furthermore, this will allow the dependency of traffic flow in a road segment with adjacent roads.

3.1 Attention Mechanism

Attention mechanisms allow the code to identify the most relevant parts of the data for prediction. In this case, the model developed by (Ashish Vaswani, 2017) was used by its Keras implementation. Specifically, this mechanism uses a multi-head attention, which projects the inputs and performs the attention calculation on each projection in parallel. These values are concatenated and projected to get the final output of the layer. This module needs three arguments: query (Q), value (V), and key (K). Its comparison with a singular head attention mechanism is depicted in Figure 3.
4. Data Collection
The methodology was tested and verified using datasets collected from Los Angeles, California. Datasets included information regarding traffic, road network, weather, touristic attractions, and Twitter.

4.1 Traffic and Road Network
The PeMS California dataset was included for traffic, using only the weekdays of May and June from 2012 in 5-minute intervals for a total of 228 sensors (Figure 4). Instead of using an adjacency matrix, and to consider co-occurrences in traffic along the road network, a correlation matrix was used, which gives the freedom of adding or reducing roads according to the computing capability available. Figure 4 shows different scenarios in which the acceptable correlation parameter was changed gradually from 0.90 to 0.25. This image also shows that, while most roads close together are correlated at a high level (i.e. 0.90) are closer together, some roads that are further away might also impact traffic on the road being evaluated.

For the models where GCN is not used, the correlation matrix is changed by the roads with highest correlation, setting the acceptable correlation parameter at 0.90. For example, when predicting traffic for road 1, the roads that achieved a correlation higher than 0.90 are roads 7, 8, 11, and 15. These roads will be added to the input matrix for non-GCN models as independent variables, having road 1 as the target variable.

4.2 Weather
The weather dataset was downloaded from the National Centers for Environmental Information, using the Los Angeles International Airport sensors. Variables used from this dataset are temperature, dew point, visibility, and wind. Since the sensors provide hourly data, the dataset was formatted to fit the traffic dataset, having 12 5-minute intervals within the hour where the weather information remained constant.

4.3 Touristic Attractions
Points of interest (POIs) were included for this study within the boundaries of Los Angeles city (Figure 4). 1142 POIs were identified and placed as a feature for their nearest traffic speed sensor. This will impact the traffic speed prediction by generating more conglomeration around these sensors.
4.4 Social Media Networks

Tweets were collected using the Twitter API for academic research v2. By using the SEARCH endpoint, we had access to the full tweet archive, where we selected all tweets posted within a 25-mile radius of the center of Los Angeles that contained words related to traffic (“i.e. traffic”, “congestion”, “traffic jam”, “delay”). 2887 tweets were collected (shown in Figure 5), which were assigned to a sparse matrix of time steps x number of roads. This will allocate the tweet in the timestep in which it was posted and to the geographically closest traffic sensor. Furthermore, sentiment analysis was applied to the tweets, which yielded a new variable ranging from -1 (negative) to 1 (positive). This was added to the variables as constant across the roads but varying in time, representing the collective social network’s mood.

Figure 5. Placement of traffic speed sensors (top), tweets (bottom left), and points of interest (bottom right) within Los Angeles City.
5. Results and Discussion
All models were run using Google Cloud Platform (GCP) c2-standard-8 machine with 8 vCPU and 32 GB memory. The train/validation split was set to 0.5/0.2 respectively, using the remainder of the data for testing. The models were run using libraries like Keras and TensorFlow, and hyperparameter tuning using Bayesian Optimization. Table 1 shows the specifications for both hyperparameters. Additionally, the specifications for the models can be found in Table 2.

Table 1: Hyperparameter settings for tuning

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min</th>
<th>Max</th>
<th>Step</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.1</td>
<td>0.0001</td>
<td>x 0.1</td>
<td>Choice</td>
</tr>
<tr>
<td>LSTM units</td>
<td>32</td>
<td>160</td>
<td>+ 16</td>
<td>Int</td>
</tr>
</tbody>
</table>

Table 2: Model Metadata

<table>
<thead>
<tr>
<th>Data</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input sequence length</td>
<td>12</td>
<td>12 timesteps prior to the target timestep were used in the LSTM stage to make the prediction.</td>
</tr>
<tr>
<td>Forecast horizon</td>
<td>3</td>
<td>The predicted forecast was 3 timesteps ahead</td>
</tr>
<tr>
<td>Acceptable correlation</td>
<td>0.9</td>
<td>Correlation parameter for adjacency matrix</td>
</tr>
<tr>
<td>Executions per trial</td>
<td>2</td>
<td>Number of executions per trials the tuner will run</td>
</tr>
<tr>
<td>Number of trials</td>
<td>3</td>
<td>Number of maximum times the tuner will run</td>
</tr>
<tr>
<td>Epochs</td>
<td>10</td>
<td>Number of times the model will run to calculate loss</td>
</tr>
<tr>
<td>Number of optimized epochs</td>
<td>100</td>
<td>Number of epochs the model will be trained for using the best hyperparameters to find the necessary number of epochs</td>
</tr>
</tbody>
</table>

From the Table 2, we infer that each model will be run a total times of executions per trial x number of trials x epochs = 60 times to find the best hyperparameters based on the loss and an additional 100 to get the epoch with the best value for validation loss.

5.1 Numerical Results
Two models were tested, GCN with LSTM and GCN with LSTM and Multiheaded Attention Mechanism. The summary of the performance of each model can be found in Table 3.

Table 3. Summary of model performance.

<table>
<thead>
<tr>
<th>Model</th>
<th>Best Hyperparameters</th>
<th>Number of best epochs</th>
<th>Validation Loss</th>
<th>Naïve MAE</th>
<th>Model MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCN-LSTM</td>
<td>Learning rate: 0.01</td>
<td>84</td>
<td>0.002457</td>
<td>0.003554</td>
<td>0.003801</td>
</tr>
<tr>
<td></td>
<td>LSTM units: 80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCN-LSTM with Attention</td>
<td>Learning rate: 0.01</td>
<td>21</td>
<td>0.002489</td>
<td>0.003554</td>
<td>0.003799</td>
</tr>
<tr>
<td></td>
<td>LSTM units: 144</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the Naïve, which are the forecasts of the last value of speed in each node, and MAE of the model we can see that neither model suffered from overfitting. The model, GCN-LSTM, shows a low validation loss and high accuracy for the testing datasets (as can be seen in Figure 6). The GCN-LSTM with Attention model also shows a low validation loss, although not lower than GCN-LSTM, is not significantly worse, and shows a slight improvement for the MAE accuracy. The number of best epochs is based on a baseline of 100, which implies that the GCN-LSTM model took a significantly longer time to train, 84 epochs, than the GCN-LSTM with Attention model.

Further experiments will be run to show lower-complexity models like LSTM and LSTM with Attention. Additionally, the question whether which model can predict with high accuracy with a partially observed traffic dataset still holds.
With the current experiments being performed with a forecast horizon of 3 5-minute timesteps, this is a short-term prediction, which would be more likely to be implemented in a real-time system. Since last-mile deliveries are in need of long-term predictions to optimize their routes before the workday, a third set of experiments with a forecast horizon of 720 5-min intervals, 12 hours, will also be run.

5.2 Graphical Results
The prediction yields forecast for each road on the adjacency matrix (228 sensors). As an example, we’ll show road #0 in Figure 6. This shows a high accuracy for both models, but the GCN-LSTM with Attention model seems to predict better in both, high-demand and low-demand traffic speed, and the GCN-LSTM model has a better prediction for low-demand.

5.3 Proposed Improvements
There are a few proposed improvements for this research: (1) Since the acquisition of Twitter on 2023 and a new CEO/owner, the Twitter API was deprecated for academic use hence this research cannot be further replicated using this social media platform. Using other media platforms like Instagram or TikTok would be a research venue worth investigating, (2) POIs have more popularity whenever people masses post about them on social media, using Instagram to generate new POIs or weights for POIs depending on the number of geotags or amount of size of network from the user posting the geotag is an ongoing effort to continue the findings on this paper, and (3) this methodology is tailored to be experimented with scarce traffic data, next steps include partially-observed speed sensor data to achieve an acceptable accuracy with the help of the other variables that are present without a strain in the transportation system’s budget.

Figure 6. Prediction (orange) and actual (blue) values for road #0 using GCN-LSTM (top) and GCN-LSTM with Multihead Attention (bottom).
6. Conclusion

After performing the literature review, some research gaps were identified. On one hand, social networks have yet to be exploited for influential data. Social media has been proposed in the literature to fill in the gap of lack of traffic data in urban areas but requires constant queries to apply data mining to sparse tweets.

On the other hand, there is still a need to find a hybrid framework that can develop a traffic prediction tailored for last-mile deliveries while using social network analysis. While methodologies and heuristics have been proposed for traffic prediction (including RNNs, CNNs, hybrid CNN-RNN models, GNN, and attention mechanisms), these are mainly tailored for developed countries. Underdeveloped countries have a data problem, either inexistent or too sparse, and there is no implementation of a multimodal architecture using social media that has been tested against partially observed traffic to simulate lack of data in non-ITS.

A hybrid methodology that can implement a social network analysis and include it in traffic prediction with multiple factors, tailored for last-mile deliveries, has been developed. By using GCN, LSTM, and Multiheaded attention mechanism, the 8-variable dataset (7 independent and 1 dependent) was processed and yielded a higher accuracy by using Bayesian Optimization. The GCN-LSTM model showed a slight improvement in validation loss, while GCN-LSTM with Attention showed improvement in training time (number of epochs) and MAE accuracy. Further research will include other social media networks, like Instagram, and further analysis of partially observed traffic to tailor the methodology for underdeveloped countries.

References


Biographies
Valeria Laynes Fiascunari is a Ph.D. student expected to graduate in Spring 2024 from the Industrial Engineering and Management Systems department at the University of Central Florida. With more than 4 years of research experience in machine learning and industrial engineering, she has participated in conferences like MIT Scale LATAM, ICAI, and NeurIPS, and has published in Applied Sciences journal. She has industry experience with resource allocation and forecasting for marketing strategies and currently works as a commercial sales executive. Her dissertation research focuses on traffic prediction for underdeveloped countries by using deep learning methodologies, but she also explores other research venues like social network analytics, discrete-event simulation, and decision support systems. Her main motivation as an Industrial Engineer is to close the gap between human-machine interactions and provide a more humanistic view on engineering solutions.

Dr. Edgar Gutierrez-Franco is a Research Affiliate at the MIT Center for Transportation and Logistics (CTL). He actively collaborates with the MIT Food and Retail Operations Lab on applied research projects, specifically focusing on retail businesses, food and agribusiness supply chain management, and circular supply chain initiatives.
Additionally, he serves as an Analytics Scientist at the International Maize and Wheat Improvement Center (CIMMYT-Mexico), where he conducts agricultural research. Before these roles, Dr. Edgar worked as a postdoctoral associate at the MIT Omnichannel Distribution Strategies Lab (2021-2022), actively participating in projects related to sustainable circular supply chains and omnichannel distribution strategies for CTL’s corporate partners. He also contributed as an instructor and advisor in the Supply Chain Management master’s program and the Graduate Certificate in Logistics & Supply Chain Management (GCLOG) program. Dr. Edgar was a visiting scholar at MIT CTL (2009, 2010), where he developed projects related to Supply Chain Innovation in Emerging Markets and Carbon-Efficient Supply Chains (LATAM and USA). His research focuses on designing and implementing data and model-driven digital twins for urban logistics, network design, and sustainable supply chain management. He also explores the intersection of applied optimization techniques with statistical and machine learning methods to support logistics and supply chain management decision-making processes.

**Dr. Luis Rabelo** is a professor at the Industrial Engineering and Management Systems Department at the University of Central Florida. He is also the Co-Director of the Simulation Interoperability Lab, which provides a collaborative computing environment that supports the creation, execution, and reuse of simulations that can integrate multidisciplinary models representing the elements of network-centric warfare. As an expert in management systems, artificial intelligence (AI), and simulation, Dr. Rabelo works with his research team to better understand the industry process, simulate deep learning design, and create next-generation systems. The group works with both public and private organizations to solve real-world problems using AI and simulation solutions designed through brain wave analysis and research. These innovations find success in a shorter period, demand more attention, and offer longer retention than traditional teaching, coaching, and employee training methods. Luis Rabelo is working on multiple-resolution modeling patents and inventions. Two of his patents have allowed neural networks to meet the requirements of civil aviation regulations and allow neural networks to fly in civilian aircraft. (US 6,577,960 and US 6,157,894). His vision for the future is becoming a reality within the walls of his simulation laboratory. With a passion for improving management systems using AI and simulation, Dr. Rabelo and his team are helping companies become more efficient, productive, and profitable. His larger goal has to do with helping the individual.