A Framework for Urban Last-Mile Delivery Traffic Forecasting: An In-Depth Review of Social Media Analytics and Deep Learning Techniques

Valeria Laynes-Fiascunari 1, Edgar Gutierrez-Franco 2, Luis Rabelo 1,*, Alfonso T. Sarmiento 3 and Gene Lee 1

1 Department of Industrial Engineering and Management Systems, University of Central Florida, Orlando, FL 32826, USA; valeria.laynes@knights.ucf.edu (V.L.-F)
2 Massachusetts Institute of Technology, Center for Transportation and Logistics, Cambridge, MA 02142, USA
3 Grupo de Investigación en Sistemas Logísticos, Facultad de Ingeniería, Campus del Puente del Común, Universidad de La Sabana, Km. 7, Autopista Norte, Chía 250001, Colombia
* Correspondence: luis.rabelo@ucf.edu

Abstract: The proliferation of e-commerce in recent years has been driven in part by the increasing ease of making purchases online and having them delivered directly to the consumer. However, these last-mile delivery logistics have become complex due to external factors (traffic, weather, etc.) affecting the delivery routes’ optimization. Intelligent Transportation Systems (ITS) also have a challenge that contributes to the need of delivery companies for traffic sensors in urban areas. The main purpose of this paper is to propose a framework that closes the gap on accurate traffic prediction tailored for last-mile delivery logistics, leveraging social media analysis along with traditional methods. This work can be divided into two stages: (1) traffic prediction, which utilizes advanced deep learning techniques such as Graph Convolutional and Long-Short Term Memory Neural Networks, as well as data from sources such as social media check-ins and Collaborative Innovation Networks (COINs); and (2) experimentation in both short- and long-term settings, examining the interactions of traffic, social media, weather, and other factors within the model. The proposed framework allows for the integration of additional analytical techniques to further enhance vehicle routing, including the use of simulation tools such as agent-based simulation, discrete-event simulation, and system dynamics.

Keywords: traffic prediction; intelligent transportation system; social media analytics; last-mile delivery

1. Introduction

Customer satisfaction determines whether a company is successful; losing a loyal customer is more expensive than trying to gain a new one. Therefore, companies have been trying to keep up with the industry, trying to innovate and give solutions to problems that escalate quickly. Specifically, with the e-commerce industry and the accelerated development of technology, people are opting to buy everything online. It has been noted that e-commerce has grown five times more than store commerce, largely due to online mobile shopping [1]. This demand pushes organizations to open an omnichannel strategy that facilitates online purchases and delivers products faster than their competitors. However, what is “customer satisfaction” in e-commerce? What does a company have to do to succeed in the industry? In the book Approximate Dynamic Programming for Dynamic Vehicle Routing [2], we find the following statement: “The settings for LSPs (Logistic Service Providers) are … changing. In particular, the expectations of shippers and customers increase. […] On the one hand, shippers and customers expect reasonably priced services. On the other hand, they expect the services to be fast and reliable. To allow reliable services and to keep the customers’ loyalty, LSPs need to consider conditions and requirements” [2].

Received: 5 April 2023
Revised: 30 April 2023
Accepted: 3 May 2023
Published: 10 May 2023
These conditions and requirements can be anything from known deterministic information such as the number of customers to visit, a list of products to deliver to unknown addresses, and stochastic information such as new orders, same-day delivery, changing traffic patterns, changing weather, etc.

A study shows that two of the most important online shopping features for users are information quality and online service quality [3]. They argue that for the business to be sustainable, they would need these characteristics embedded in policies, strategies, and on their website.

This research is backed up by an investigation where the purpose was to develop the drivers and outcomes of online shopping experiences [4]. The authors found that the principal driver of the online shopping experience is product/service experience quality, which includes delivery as the most important aspect of this category. This driver even comes before retail price. This means that customers prefer how the product is delivered and the customer service over the price of the product they buy. It seems trivial if we portray ourselves as the ultimate customer when we know that we have a package arriving on a certain day and keep checking the tracking history and updates nonstop. However, if the package does not arrive on the predicted date, our customer satisfaction decreases.

The industry grows fast for B2B and B2C e-commerce companies, and management faces a big challenge: the bar keeps rising, 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 that usually leads to poor vehicle utilization and service duplication [5]. Therefore, they set a list of challenges that last-mile delivery services face, such as (1) demand patterns and 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, such as distribution centers, budget, vehicles, drivers, etc. to fulfill the demand rate.

Even if traffic can be predicted from weather and historical traffic data, it does not have the desired accuracy for more residential areas. In addition, ITS still faces a budget challenge: traffic sensors are not widely installed in urban areas due to their cost, and are mostly found on highways and freeways, where most accidents happen due to the increased speed limits. Therefore, logistic companies still need a framework that will not disrupt ITS planned resources and can be used for effective traffic prediction regardless of the area of delivery.

Currently, cities use traffic sensors mostly on highways or freeways due to lack of budget. As a result, research has been predicting traffic to prevent accidents on these types of roads or to have a better infrastructure for the traffic management system. However, shifts in e-commerce and logistics demand more from Intelligent Transportation Systems (ITS), which are a challenge to implement, and this calls for resources such as money, workforce, and time.

This issue leads to inaccurate, sometimes non-existent, traffic prediction for urban and residential areas, where most packages are delivered to end customers (nano stores, houses, residential buildings, communities, etc.). There is a potential solution using social media, which hubs millions of geotagged data regarding check-ins, events, and movement from masses of people. While smart cities have developments in areas such as smart grids and autonomous vehicles [6,7], ITS still lacks a framework where social media data, which has been applied to supply chain management and e-commerce [8], is implemented in a traffic prediction tailored for last-mile deliveries. The need for this framework is derived from companies such as Amazon, where optimal routes are calculated and given to drivers before starting the day. However, drivers often change them because the ones previously retrieved are not optimal given the circumstances at the time of delivery.
Dynamic Vehicle Routing Problems (DVRP) are a must for the supply chain, given that traffic is inconsistent throughout the week or the day, and these changes may cause delivery delays. The same route used for one day might not be optimal for the next one, not even for the same customer list delivered during a different shift. Having a framework to optimize DVRPs without demanding more from ITS by using external factors now looks like the next step to optimize last-mile delivery.

The following research questions were derived before conducting the systematic literature review and considering the challenges stated in the Problem Statement. These questions are the basis of this research. In addition, these questions are the main part of the eligibility criteria for the deletion and inclusion of papers in the study and further analysis.

(a). Is it possible to characterize an urban road network with dynamic data, such as weather, traffic, and social media updates, to implement into a Dynamic Vehicle Routing Problem methodology to improve last-mile delivery?

(b). Can a local social network’s behavior contribute to traffic prediction?

(c). Could this framework be an initial point to further implement real-time route optimization, if needed, for highly variable optimal routes?

The primary contributions of this research can be summarized as follows: (1) a systematic literature review following the PRISMA framework to summarize the contributions conducted in social media usage for traffic prediction and traffic usage for vehicle routing; (2) an illustrative example of the traffic prediction methodology and a proposal for a new social network feature extraction.

In Section 2, we conducted a literature review of existing research on traffic prediction and social network analytics to provide context for our study. Section 3 describes the methodology used in our research, including the data collection and analysis techniques applied. Section 4 presents the framework for traffic prediction that we propose, which is based on social network analytics. In Section 5, we provide an illustrative example to demonstrate the effectiveness of our proposed framework. Finally, in Section 6, we present our conclusions and discuss future work in this area.

2. Literature Review and Conceptual Context

This literature review was conducted following the PRISMA framework [9]. No previous research protocol was followed to conduct this literature review. The search strategy focused on the existing solutions to last-mile dynamic VRPs, with special attention to solutions including deep learning and weather/social media factors influencing traffic conditions. The search only considered peer review journal articles and conference papers from Compendex/Inspect, Web of Science, and IEE that were published during the last six years (2014–2020) and the ones that will be published in 2021. The following combination of topics was searched in the title, keywords, and/or abstract: Weather AND Traffic, Weather AND Route, Weather AND Last-Mile, Route AND Traffic, Last-mile AND Route, Last-mile AND Machine Learning, Machine Learning AND Route, Weather AND Traffic AND Route, Last-mile AND Route AND Traffic, Last-mile AND Route AND Machine Learning, Route AND Deep Learning AND Weather, Social Media AND Traffic, Social Media AND Traffic, Social Media AND Route, Social Media AND Traffic, Social Media AND Route. The operator OR was used to group all the keywords from a topic. A detailed list of the topics with their respective keywords is shown in Table 1.

Some topics were excluded during the exploration phase, such as optics, telecommunication networks, ship routing, offshore delivery, air traffic, air control, air pollution control, unnamed aerial vehicles, power grids, public transportation, road safety, bike trips, etc. The PRISMA framework for systematic search steps was conducted. A summary of the steps is shown in Figure 1.
Table 1. List of topics and respective keywords.

<table>
<thead>
<tr>
<th>Weather</th>
<th>Last-Mile</th>
<th>Route</th>
<th>Machine Learning</th>
<th>Traffic</th>
<th>Social Media</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather</td>
<td>Last-Mile</td>
<td>Routing</td>
<td>Neural networks</td>
<td>Traffic flow</td>
<td>Social media</td>
</tr>
<tr>
<td>Climate</td>
<td>Last-Mile</td>
<td>Route</td>
<td>Deep learning</td>
<td>Traffic volume</td>
<td>Social networks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Routes</td>
<td>Reinforcement learning</td>
<td>Traffic speed</td>
<td>Geotag</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vehicle routing</td>
<td>Neural network</td>
<td>Traffic prediction</td>
<td>Geolocation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vehicles routing</td>
<td>Expert systems</td>
<td>Traffic</td>
<td>Location-based</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urban routing</td>
<td>Expert system</td>
<td>Traffic congestion</td>
<td>Location based</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Machine learning</td>
<td>Transit</td>
<td>Twitter</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Data mining</td>
<td>Traffic jam</td>
<td>Instagram</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Google places</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Facebook</td>
</tr>
</tbody>
</table>

Some topics were excluded during the exploration phase, such as optics, telecommunication networks, ship routing, offshore delivery, air traffic, air control, air pollution control, unmanned aerial vehicles, power grids, public transportation, road safety, bike trips, etc. The PRISMA framework for systematic search steps was conducted. A summary of the steps is shown in Figure 1.

Figure 1. PRISMA framework for systematic search steps. Source: [9].

Using the steps and framework from PRISMA guidelines, 4957 papers were selected during the Identification phase. After duplicate removal, there were 2370 papers. Next, during the Screening phase, titles and abstracts were evaluated, and 1837 papers were excluded because they included one or more of the following topics: traffic accident analysis, parking problems, ship routing, speed analysis (without considering traffic), traffic psychology, fiber/optical networks, air traffic, road conditions, the crowd flows, pollution analysis, data collection methods, electric vehicle, unmanned aerial vehicles, wireless networks, humanitarian logistics, drone deliveries, traffic assignment problem, green routing, traffic flow distribution, emergency routing, parking search, arc routing problems, inventory problems, aviation, and VANETs. After discarding papers, 137 articles were assessed for eligibility, and 396 were excluded due to eligibility criteria stated in
previous steps or were not aligned with research questions. Finally, 137 articles were analyzed, and 56 were included in the literature review.

Figure 2 presents a co-occurrence word graph visually representing the relationships between key terms in traffic prediction, urban logistics, and vehicle routing vocabulary. The graph was created by analyzing a corpus of relevant literature in these fields and identifying the terms that frequently appear near each other.

![Figure 2. Relationships in Traffic Prediction, Urban Logistics, and Vehicle Routing with Co-occurrence of Literature Review Graphs.](image)

The main clusters in the graph represent groups of terms that are closely related to each other. For example, the first cluster, represented by the red color, is labeled “vehicle routing and heuristic algorithms”. This cluster includes terms such as “vehicle routing”, “heuristic”, “algorithms”, and “optimization”. These terms are closely related to finding the most efficient vehicle routes in an urban environment.

The second cluster, represented by the blue color, is labeled “traffic congestion”. This cluster includes terms such as “congestion”, “traffic flow”, “gridlock”, and “bottlenecks”. These terms are closely related because they pertain to the negative effects of excessive vehicle traffic on urban roadways.

The third cluster, represented by the purple color, is labeled “traffic engineering and intelligent transportation systems”. This cluster includes terms such as “traffic engineering”, “intelligent transportation systems”, “traffic management”, and “transportation planning”. These terms are closely related because they pertain to the application of technology and engineering principles to improve the efficiency and safety of urban transportation systems.

The fourth cluster, represented by the green color, is labeled “transportation and machine learning”. This cluster includes terms such as “machine learning”, “artificial intelligence”, “transportation”, and “prediction”. These terms are closely related because they pertain to the application of machine learning techniques to improve the efficiency and safety of urban transportation systems.

Overall, the co-occurrence word graph in Figure 2 provides a good visual representation of the relationships between key terms in traffic prediction, urban logistics, and vehicle routing vocabulary. It helps us to understand the relationship between the different terms.
and how they are interconnected. It also helps us to identify the main areas of research in this field and the main contributions in these areas.

2.1. Traffic Prediction

Traffic prediction is now widely used in cities for infrastructure planning, avoiding congestion, and proactive measures in case of accidents. However, due to its high complexity and the fact that it is affected by external factors, a parametric model is insufficient for acceptable accuracy. Therefore, many nonparametric models have been used for traffic classification and/or regression, such as tree algorithms, different neural networks, and K nearest neighbors. The studies presented in this section use two different data collection techniques (floating car or sensor), two different types of prediction (classification or regression), and two different types of areas (urban or highway/freeway). A summary of these characteristics can be found in Table 2 at the end of this section. Additionally, there has been an increasing trend in traffic prediction studies, as shown in Figure 3, which shows the importance of this research and the search for a higher-accuracy model that can correctly represent real-life scenarios.

![Number of Traffic Prediction Studies](image)

**Figure 3.** Traffic prediction studies throughout the years.

Long-Short Term Memory (LSTM) neural network is a recurrent network and handles time series using prior timeframes. LSTM is the most popular method. For example, Huang et al. [10] used LSTM to predict traffic congestion to help drivers with their routes. The congestion was determined using real-time images gathered by drones. The drones were able to obtain information regarding speed and volume. LSTM was also used for forecasting using a floating car dataset collected from sensors [11]. The scheme was compared to Support Vector Regression (SVR). LSTM showed better robustness and accuracy. LSTM also has instances, such as [12], that predict multi-step traffic and capture long- and short-term dependencies.

Different techniques can be used but their effectiveness is limited, such as the perceptron with several hidden layers (MLP), recurrent networks (RNN), and regression and classification trees. For example, we started with an MLP designed using batches [13]. However, this study was limited from the viewpoint of external variables and other situations. The sub-predictor fusion was achieved with Bayesian models, avoiding errors of relatively larger magnitudes. However, deep learning is still used with RNNs to make advanced regression models, and thus be able to predict traffic on the roads. A tree model was effectively used, and so was the Gradient Boosted Regression method [14]. The problem is that these methods are not applicable in urban areas since the sensors have limitations.

More specific models using urban settings are also found in the literature. For example, two studies using urban and road/highway sensor data for training an LSTM NN model [15,16] take advantage of the LSTM’s sequence-to-sequence (seq2seq) feature. In addition, these studies emphasize the use of traffic-related characteristics to predict, such as immediate traffic lanes and correlations of prior periods with the existing forecast.
Important traffic characteristics are included in network schemes that consider the spatial relationships of different roads. For example, Graph Convolutional Networks (GCNs) exploit the correspondence between a graph of multiple dimensions and a network of roads. Converting the road network into a road graph and its connectivity in a seq2seq model is a straightforward application [17]. First, seven categories of roads were devised to classify the roads. This categorization was then used as an additional input. More sophisticated systems use GCN, transforming forecasted positions into nodes and contiguosness to other roads as edges, and then converting this model into the input for an LSTM NN, which can be used for short-term projections [18]. This research concluded that graph size does not determine the model’s performance. Therefore, GCN can take parts of urban areas and perform specific predictions.

We have discussed research that focuses on traffic-related datasets. Therefore, all studies presented up until this section only consider traffic-related datasets (e.g., dates, timestamps, historical data, holidays, etc.). Nevertheless, as previously stated, an accurate and robust prediction should include more factors if it highly influences traffic. Therefore, the research conducted in this area is presented below.

2.1.1. Using Weather

More realistic models include the weather variable. Non-deep learning methodologies in this section use K-Nearest Neighbors with a trend adjustment considering past traffic flows to predict a future state [19] or to dynamically adjust parameters according to weather conditions [20]. Random forest is a common out-of-the-bag classification algorithm used to predict traffic states using a congestion index based on historical data [21]. It could also be used in regression problems [22], where weather and other spatiotemporal features are available.

Recursive Least Square [23] was used for average speed prediction in urban areas with floating cars data. One first instance used Kalman Filter as the prediction model while using RLS to merge average speed with other factors [24]. This study showed that there is a high correlation between the six-time slots previous to the predicted time slot. Another paper addressed the long-term prediction issue by using RLS for parameter estimation and multi-scale correlation for prediction by categorizing weather into categories according to humidity.

Two other non-deep learning algorithms include Bayesian Networks for a probabilistic approach to traffic congestion [25] and Support Vector Regression with a hybrid evolutionary algorithm implemented with a genetic algorithm to predict traffic flow [26]. However, neither of these methods have proven to have a higher accuracy than neural networks in research. Moreover, traffic prediction requires advanced feature computations since it is a highly complex and non-linear problem that requires advanced feature computations.

Three deep-learning studies were performed in highway or freeway settings. Traffic prediction was more accurate than SVR, GBRT, and LSTM [27] using a traditional backpropagation-based neural network. It used a layer per term, defining the long term as traffic flow, the medium term as driver’s travel habits, and the short term as the numeric traffic flow variations. However, it was not used in an urban setting. Furthermore, it used 36 levels for the weather variable, which questions the learning rate and accuracy of the model, which may be seen as overfitting the dataset. Another article used MLP and mutual information, based on the K value from KNN, to identify the interdependency of historical traffic and future traffic [28]. What made this study relevant is not the use of MLP, which is outdated and less reliable than LSTM. Still, the model can detect inaccurate, noisy, and faulty data—common traits of datasets collected by traffic sensors.

The third study from this list uses a recurrent neural network and a gated recurrent unit (GRU) [29]. This method discovered non-linear temporal correlations between weather variables and traffic flow but failed to apply the framework to smaller roads. Finally, a regression algorithm starting with different clustering roads using the Gaussian Mixture Model (GMM) was applied in an urban setting and predicted using an artificial neural
network [30]. Furthermore, it can perform adaptive forecasting, but it must manually be updated whenever lane reductions occur.

LSTM NN is the most adopted. LSTM NN has outperformed Deep Belief NNs using precipitation as the weather variable [31]. LSTM NN had better accuracy than decision trees, SVM, and MLP using online traffic and weather data [32]. Rainfall and temperature improve traffic prediction accuracy [33]. A novel application of a model used for traffic prediction without weather variables [18], presented in the previous section, used linear graph transformation to set roads as nodes and adjacency [34], as shown in Figure 4. This architecture took advantage of the GCN trait of being a multidimensional model.

![GCN-LSTM System Architecture](based on [34]).

**Figure 4.** GCN-LSTM System Architecture (based on [34]).

### 2.1.2. Using Social Media

Traffic forecasting methodologies using social networks address the problem of inaccurate urban forecasts, especially where there is limited availability. The “human” sensors use Twitter as a hub. An application was developed to create a traffic network on Twitter using spatial features. Thus, the problem of the scarcity of data was overcome [35]. This technique, called Traffic Prediction (CTP), can predict traffic without historical data. The correlation between traffic conditions and tweets [36] was investigated, but the results indicate that systems based on historical data are superior. T-MAPS was an application created to imitate the Routing of Google Maps. Its operation has the following steps: segmentation to a personalized section of the city, creating static networks with a timeline, and applying weights to the network’s edges based on the traffic.

On the other hand, these networks do not merge with the dataset’s features. This merging could be performed with a GCN; therefore, T-MAPS routes achieved 60% accuracy compared to Google Maps. However, it is the first traffic prediction model that uses social networks and is designed for route generation.

Rarely does the traffic forecast need social networks to be more accurate. However, as accidents increase, some studies use event detection for real-time and dynamic forecasting using tweets and the weather [37,38]. This research incorporated a full set of variables (social media, timestamp, weather, and traffic information) influencing traffic congestion. This makes such studies unique, and they are examples of good predictions.

GCN and LSTM embedded social media using queries and travel analysis [39]. Users searched for routes, points of interest, and travel times. These searches affected traffic flow. The model predicted traffic in an urban environment using a seq2seq. Therefore, the model required a dataset that included traffic flows, queries, and road networks.

These studies are applicable to increasing the accuracy and reducing the delivery time of routes in a vehicle routing problem. However, frameworks incorporating social media in routing problems are for touristic purposes and visiting points of interest (POIs) [40]. Even though analyzing POIs in road networks can benefit last-mile deliveries, no known research includes social media in a situation-aware route or last-mile methodology.
### Table 2. Traffic prediction articles and characteristics.

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>Variables</th>
<th>Prediction Type</th>
<th>Area</th>
<th>Method</th>
<th>Traffic Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>[38]</td>
<td>Weather, social media</td>
<td>Regression</td>
<td>Highway</td>
<td>Bayesian Networks</td>
<td>Sensor</td>
</tr>
<tr>
<td>2016</td>
<td>[37]</td>
<td>Weather, social media</td>
<td>Regression</td>
<td>Freeway</td>
<td>Deep Belief NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2016</td>
<td>[32]</td>
<td>Weather</td>
<td>Classification</td>
<td>Freeway</td>
<td>LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2017</td>
<td>[29]</td>
<td>Weather</td>
<td>Regression</td>
<td>Freeway</td>
<td>GRU NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2018</td>
<td>[27]</td>
<td>Weather</td>
<td>Regression</td>
<td>Freeway</td>
<td>Cascaded NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2018</td>
<td>[39]</td>
<td>Social media</td>
<td>Regression</td>
<td>Urban</td>
<td>GCN NN, LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2019</td>
<td>[34]</td>
<td>Weather</td>
<td>Regression</td>
<td>Urban</td>
<td>GC NN, LSTM NN</td>
<td>Floating car</td>
</tr>
<tr>
<td>2019</td>
<td>[33]</td>
<td>Weather</td>
<td>Regression</td>
<td>Freeway</td>
<td>LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2020</td>
<td>[16]</td>
<td>-</td>
<td>Regression</td>
<td>Urban/highway</td>
<td>LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2020</td>
<td>[18]</td>
<td>-</td>
<td>Regression</td>
<td>Freeway</td>
<td>GCN NN, LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2020</td>
<td>[17]</td>
<td>-</td>
<td>Classification</td>
<td>Urban/highway</td>
<td>GNN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2020</td>
<td>[10]</td>
<td>-</td>
<td>Classification</td>
<td>Highway</td>
<td>LSTM NN</td>
<td>Sensor</td>
</tr>
<tr>
<td>2022</td>
<td>[12]</td>
<td>-</td>
<td>Regression</td>
<td>Freeway</td>
<td>LSTM NN</td>
<td>Sensor</td>
</tr>
</tbody>
</table>

#### 2.2. Situation-Aware Route Planning

Situation-aware route planning entails more variables than just the vehicles, the starting position, and the destination. It is usually studied for problems that include the vehicle’s initial position and destination. Still, it could be applied at each delivery during the last mile to improve the estimated arrival time [41]. Dynamic, sometimes stochastic, variables such as traffic and weather come into play. However, many data collection methods are unreliable and can lead to high sparsity [42].

On the other hand, using these datasets to route vehicles in real time creates agglomeration on previously empty roads. This section presents several studies that have investigated dynamic routing phenomena. A summary of these studies can be found in Table 3.
### Table 3. Situation-aware routing planning articles and characteristics.

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>Variables</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>[43]</td>
<td>Traffic</td>
<td>Dijkstra</td>
</tr>
<tr>
<td>2016</td>
<td>[45]</td>
<td>Traffic</td>
<td>Dijkstra</td>
</tr>
<tr>
<td>2016</td>
<td>[46]</td>
<td>Weather</td>
<td>RSA, CEPO</td>
</tr>
<tr>
<td>2017</td>
<td>[47]</td>
<td>Traffic</td>
<td>Ant Colony, Dynamic Graph</td>
</tr>
<tr>
<td>2017</td>
<td>[48]</td>
<td>Traffic</td>
<td>Ant Colony</td>
</tr>
<tr>
<td>2018</td>
<td>[49]</td>
<td>Traffic</td>
<td>Floyd’s, DBSCAN</td>
</tr>
<tr>
<td>2018</td>
<td>[50]</td>
<td>Weather</td>
<td>RSA, CEPO</td>
</tr>
<tr>
<td>2018</td>
<td>[51]</td>
<td>Traffic, weather</td>
<td>DB NN</td>
</tr>
<tr>
<td>2019</td>
<td>[52]</td>
<td>Traffic, Weather, Distance</td>
<td>Autoencoder NN</td>
</tr>
<tr>
<td>2019</td>
<td>[53]</td>
<td>Traffic</td>
<td>k-shortest path, LSTM NN</td>
</tr>
<tr>
<td>2019</td>
<td>[55]</td>
<td>Traffic</td>
<td>LSTM NN</td>
</tr>
</tbody>
</table>

#### 2.2.1. Routing Using Traffic Conditions

Vehicle routing using traffic can be classified by the usage of the data: real-time usage, prediction usage, or hybrid. Moreover, this data can come from experienced drivers such as taxicabs or sensors on specific roads. Two studies used the Ant Colony Algorithm (ACO) to find the best possible routes. The first one started its framework by predicting traffic and using real-time data only for non-recurrent heavy congestion [48]. However, the principal objective was not to look for a set of routes but to use ACO and jam time windows for each road to make updates to the route every time the car was positioned in a junction node of the traffic network. The second study used a set of three dynamic network graphs: distance, speed, and travel time, when considering traffic [47]. It used ACO for vehicle routing, and it updated the routes with the depreciation of travel time according to the dynamic traffic graph.

A common method to find the best paths is the Dijkstra algorithm. A hybrid model predicted short-term near-future traffic and minimal real-time information to feed the Dijkstra algorithm, outperforming static and dynamic routing [45]. A real-time usage model performed two different methods. The first one was based on the traffic flow propagation in the network using spatiotemporal correlation. In contrast, the second was based on the time-varied spare flow capacity of the road link in question [43]. This model updates traffic flow input data every time interval, showing that proactive route guidance leads to fewer re-routings for each rerouted vehicle.

A more advanced hybrid model using another simple routing algorithm, k-shortest paths, selects a pre-set pool of possible and optimal routings and predicts future traffic states with LSTM NN. This approach does not require a constant flow of real-time information [53]. However, it requires probe cars to detect non-recurrent congestion and uses this information to eliminate K-shortest paths that contain congested roads. Another study used LSTM to predict traffic but incorporated a double-rewarded value iteration network (VIN) to make the routes [55]. The VIN was fed the traffic prediction and data from a taxicab driver dataset, incorporating experience and knowledge of traffic trends. The proposed model achieved human-like responses, but for optimal vehicle routing, a human-like response is not optimal given the dynamic variables. Finally, a deep learning framework used Deep Belief Neural Network Classifier and historical experience data merged with weather variables to identify traffic congestion features and make a route selection [51].

Other types of machine learning tool can be used for traffic rerouting, such as DBSCAN, a density-based type of clustering. The framework that has used this tool has three steps:
(1) differentiate populated roads from non-populated roads, (2) assign weights to routes based on the traffic intensity at a given point in time, and (3) suggest routes based on user’s input [49]. This framework develops a traffic density factor using sensor data and previous traffic density factors. However, since it relies solely on sensor data, it is prone to malfunctions.

One of the previous studies used the A* algorithm for vehicle routes based on the spatiotemporal correlation of roads in a traffic network [54]. It used real-time sampling data to compute a congestion coefficient from the number of state vectors sent by monitored cars.

2.2.2. Routing Using Weather

There are a few applications of situation-aware routing using weather. However, an advanced Agent-Based Model (Q.C. Lu, Zhang, Peng, and Rahman) uses multiple agents that include alterations to a road-traffic network [44]. There are four different agents in this model: (1) traffic load, (2) weather condition, (3) type of road, and (4) travel estimation. While the first three agents perform updates on the network, the travel estimation agent runs an improved ripple-spreading algorithm (RSA) to generate the routing choices.

Another study uses the same routing algorithm offline, where the optimization is completed beforehand, and the weather is included in a coevolutionary algorithm [46]. One single routing optimization introduces the coevolutionary path optimization (CEPO) that simulates dynamic weather instead of static conditions. An upgrade on this study identifies the K-shortest paths [50] in a given dynamic environment without needing a memory-extensive hypergraph.

A recent study used an Autoencoder Neural Network to promptly predict traffic conditions, considering the possible application of this method in logistics distribution [52]. The authors believe that traditional algorithms usually cannot predict extreme conditions for the logistics’ environment, which yields a low accuracy for path optimization. Instead, they use variables such as historical traffic, weather, and distance of the road segment to make predictions.

All these methodologies consider vehicle routing from its current position to the next but do not plan, since the models are not fed a set of destinations. In this case, applying some frameworks to last-mile delivery would need an improvement that can yield a set of routes per destination.

2.2.3. Dynamic Vehicle Routing Problem

Most dynamic vehicle routing problems (DVRP) include traffic updates, since it is a complex dynamic factor and can have non-recurrent scenarios, several stochastic incoming customer orders, or a combination of both. Including stochasticity in a VRP can yield optimized vehicle routes when presented with new information. On the other hand, it is beneficial to the supply chain’s stakeholders to have a planned set of routes beforehand. DVRPs present ambiguous benefits, which depend on the situation, resulting in heuristics specific to the problem. These solutions are presented in this section, and a summary can be found in Table 4.

New customer orders are common in recent DVRPs. However, a study implemented order cancelations as well by using the Hybrid Neighborhood Search (HNS) heuristic and the Variable Neighborhood Search (VNS) to refine the solution space [56]. Since the methodology accepted changes in the customer list throughout the day, heuristics were used to insert information online and to perform a re-optimization while the vehicle was en route. In addition, fuzzy time windows were added as the quantification of the time window interval studied a constraint and service level. The short computational times show that the algorithm is appropriate for real-time scenarios in logistics.
Table 4. Dynamic Vehicle Routing Problem articles and characteristics.

<table>
<thead>
<tr>
<th>Year</th>
<th>Reference</th>
<th>Routing and Optimization</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>[56]</td>
<td>Heuristics, Hybrid Neighborhood Search</td>
<td>Fuzzy time windows</td>
</tr>
<tr>
<td>2015</td>
<td>[57]</td>
<td>Ant Colony, Local Search</td>
<td>Time windows, capacitated, heterogeneous fleet, split delivery</td>
</tr>
<tr>
<td>2016</td>
<td>[58]</td>
<td>ADP, Heuristics</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>[59]</td>
<td>Particle Swarm Optimization</td>
<td>Time windows, new demand urgency</td>
</tr>
<tr>
<td>2017</td>
<td>[60]</td>
<td>Local Search</td>
<td>Time windows</td>
</tr>
<tr>
<td>2017</td>
<td>[61]</td>
<td>Ant Colony Optimization</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>[62]</td>
<td>Improved Simulated Annealing</td>
<td>Fuel consumption</td>
</tr>
<tr>
<td>2017</td>
<td>[63]</td>
<td>Wolf Pack Algorithm</td>
<td>Fuzzy time windows</td>
</tr>
<tr>
<td>2018</td>
<td>[64]</td>
<td>RNN, Reinforcement Learning</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[65]</td>
<td>Genetic Algorithm</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>[66]</td>
<td>Evolutionary Algorithm</td>
<td></td>
</tr>
</tbody>
</table>

Another study also highlighted the importance of service level as responsiveness [57]. The objective was to optimize the responsiveness, not the total distance of the vehicle. The authors restarted the optimizer, Ant Colony Optimization, when new information was added to the problem. The insertion procedure only started when the ACO provided a significantly longer solution. The improvement on the ACO algorithm was a pheromone conservation tactic, where valuable information is passed on from one optimization to the next (i.e., the attractiveness of a path is not erased in the following matrix but stored, which diminishes the optimization time). This study yielded a fast heuristic for changes made en route, considering constraints such as time windows, heterogeneous capacitated fleet, and split delivery. Still, the scene where it was applied was too small to be applicable in a real-life situation. The effectiveness of this study in a large scenario is unknown.

Agent-Based Modeling was also used to develop a simulation with two types of agents (trucks and retailers) interacting with each other when new customer information was available [59]. Truck agents performed route optimization using Particle Swarm Optimization, depending on retailers’ new demands and urgency levels. Retailers, on the other hand, had time windows and a three-degree lateness tolerance that defined the urgency level. This study showed by simulation that the higher degree of dynamism in the routes, the longer the distance traveled by the vehicle. Furthermore, it showed that ABM is an appropriate tool to validate the environment in which different stakeholders operate for last-mile delivery.

Customer satisfaction is another common objective for DVRP. An article studied the total cost of distribution using an improved Wolf Pack Algorithm, which uses the characteristic of cooperative hunting of wolves [63]. This improvement increases the communication between artificial wolves, which results in a higher global knowledge and exploration ability. In addition, they considered dynamic changes in traffic flow and fuzzy time windows. This improved algorithm uses less computation time due to the reduced number of iterations needed and high efficiency, converging into the optimal global solution in less time than the genetic algorithm.

Instead of fuzzy time windows, regular time windows are not as easily related to customer satisfaction or service level. However, research still includes them in DVRPs. A framework included this constraint in a DVRP with travel time uncertainty due to
traffic [60]. During the design stage of the routes, the service was partitioned into areas according to an intimacy degree. This design calculated the percentage of scenarios in which a pair of customers were grouped into the same cluster. Local search was used to improve the quality of customer clusters. Vehicle routes were different in every traffic condition scenario. In other words, it was not an in-route optimization but, rather, a set of solutions that had been obtained before the workday. This approach did not consider, however, that traffic is a continuous variable, which thus has a limited number of scenarios. Therefore, it is not generalized to non-occurrent traffic states. Additionally, delivery service time was not included in the solution, and it would potentially change the routing depending on the time spent and traffic conditions by the end of the service.

One of the most recent time window constraint applications in DVRP used the time of arrival as the objective function for optimization [67]. The authors used an improved genetic algorithm to set an initial group of routes and optimized them using a hill-climbing algorithm. Data included in this optimization were road conditions, floating car data, and non-occurrent events. In addition, the network graph was modified using traffic data in intervals. Nonetheless, there was no practical application of this methodology.

Implementing constraints in DVRP increases the computational complexity, and this is a challenge if the optimization is conducted online and when the vehicles are executing the routes. Studies have tried to include time windows to increase customer satisfaction and service level, which is the most important constraint for DVRPs due to them being a responsive problem, contrary to static VRPs. Fuel consumption could be another constraint, such as the problem analyzed in a study that used stochastic speeds in the road network to simulate actual traffic [62]. However, the optimization was conducted offline, and the routes are not updated. Other types of research simplify the DVRP into a VRP, called a time-probabilistic VRP [61], where they use common methodologies (for instance, ACO) for route optimization.

Research concentrating on the DVRP without constraints usually improves the architecture to diminish online computing time. For example, a study used approximate dynamic programming (ADP) to avoid the curse of dimensionality when dealing with a DVRP with traffic congestion [58]. They used a non-stationary normal distribution to include traffic on roads and heuristics with a rollout algorithm to optimize the problem. In another instance, modifying an RNN using only the decoder significantly diminished the computational complications without sacrificing efficiency [64]. This later study states that since the VRP is not a seq2seq problem, it should not use the encoder from the RNN. Finally, they finalized the framework by using reinforcement learning for route optimization.

Evolutionary algorithms usually take long periods to finish the computations. Even if it is extensive and not optimal to include them in DVRPs, some researchers found a couple of ways to apply them due to their efficacy in solving the problem. Running the optimization in the GPU is one of the choices [65]. The authors implemented a DVRP with new customer orders throughout the day and solved it through periodic re-optimization. In other words, they were solving a classic VRP at each period, which the user defines. The results showed that running a genetic algorithm in the GPU was faster than solving the same algorithm in the CPU or an ACO.

Other application includes a self-adaptive evolutionary algorithm [66]. The article states that solutions scattered over the search space can be captured better by the dynamic changes of the VRP. The changes were modeled after generating random numbers for traffic factor in edges at every time interval. The solutions for the oncoming changes are not from scratch but an evolution of configurations from the previous solution. This way, they can inherit problem-specific knowledge from their respective parent solutions. The new configurations are encoded into the new DVRP routes. The search process for the new solution can then use different configurations to effectively handle the dynamic changes and guide the search to a promising optimal.
2.3. Summary and Research Gap

After performing the literature review, some research gaps were identified. Methodologies and heuristics have been proposed for traffic prediction and optimal vehicle routing. However, there is still a need to find an inclusive framework to develop a traffic prediction tailored for last-mile deliveries. DVRPs have been concentrating on optimizing a real-time online problem to satisfy responsiveness. Still, there is no implementation of a global methodology that can help improve it while implementing an accurate traffic prediction for urban areas with limited data available from local ITSs.

Social media has been proposed in the literature to fill in the lack of traffic data in urban areas, but it requires constant queries to apply data mining to sparse tweets. All the social media applications for traffic prediction that have been analyzed have not included other factors. Furthermore, a social network analysis is a potential solution to eliminate constant queries and find influential accounts that drive changes in traffic conditions.

A methodology that can implement a social network analysis and include that in traffic prediction with multiple factors tailored for a DVRP has yet to be developed. However, machine learning is an effective way to build this traffic prediction model, using deep learning techniques to handle complex structures.

3. Research Methodology

The research methodology describes the steps to achieve a better understanding and accuracy for last-mile deliveries under dynamic and real circumstances. Figure 5 shows the summary of the research methodology and an overview of the framework.

Figure 5. Research methodology.

The research initiated with a problem definition and a hypothesis on the existence of a framework that can characterize urban roads for last-mile delivery and what establishes the similarities between theoretically optimized routes and routes decided by drivers. In addition, a literature review was performed. This aspect was discussed in Sections 1 and 2.

Some gaps were identified after the literature review. Firstly, traffic management lacks a complete framework that includes a set of variables that can accurately yield a prediction. Research has shown that weather and social media affect traffic variability, but Intelligent
Transportation Systems still lack the inclusion of all these factors along with contextual ones. Social media inclusion in traffic prediction requires constant queries, which can be computationally extensive and too demanding. In addition, the identification of important traffic influencers is needed to avoid expensive queries from the API, given that small social media accounts do not affect traffic variability in a significant way for logistics purposes. Finally, logistic companies are still trying to find a way to develop a solution for a real-life dynamic vehicle routing problem that can yield optimized routes with high responsiveness and can imitate drivers’ expertise and tacit knowledge.

4. Social Media-Driven Traffic Prediction Framework with Illustrative Example

This methodology has four main processes: historical data collection, data analysis and prediction, dynamic vehicle routing, and validation. The following sections detail each sub-process and the proposed steps to execute the steps.

4.1. Historical Data Collection

Since floating car data is extensive to collect, sensor readings will be used for traffic data. This dataset is widely used in literature, and cleaning tools are widely studied since sparsity is a problem for urban sensors. Data cleaning and preparation will depend on the transportation systems of the cities and the availability of sensors.

The weather dataset will be collected from the National Centers for Environmental Information (https://www.ncei.noaa.gov). It manages one of the world’s largest archives of weather sensors, can offer multiple sensor locations, contains more than 37 petabytes of data, and is public for research. Variables such as cloud cover fraction, temperature, humidity, precipitation, pressure, snowfall, and snow depth, among others, are included in the datasets.

Social media data will be collected from check-in datasets from Twitter, Foursquare, or other location-based social networks. This process will derive user updates and the social network’s behavior. It will further be implemented in the graph convolutional network to characterize road segments.

4.2. Data Analysis and Prediction
4.2.1. Social Network Analysis

Research conducted on traffic prediction using social media is scarce. It either uses social media and historical traffic data [35,36,39] or combines these with the weather [37,38]. However, data mining methods to identify tweets for traffic information can be extensive and resource demanding, especially if there is a constant inflow of Tweets to be analyzed. A social network analysis with Collaborative Innovation Networks (COINs) is proposed to solve this problem.

COIN is a term developed by Peter Gloor, a researcher from MIT Sloan’s Center for Collective Intelligence [68,69], and defines a team of collaborative innovation that uses tools such as email, social networks, chat, blogs, etc., to promote a trend. People collaborating in a COIN are intrinsically motivated, have a shared vision, collaborate through the Web in teams, and work for the same goal. In addition, COINs have a common trendsetter. For example, Figure 6 shows a COIN of technology tech giants (Apple and Samsung) and the respective trendsetters. Trendsetters in a COIN are leaders who have followers and have a big impact on the network.

For this work, trendsetters will be called “traffic influencers”. Traffic influencers will have a factor determining the differential change/influence they cause in traffic conditions. Hence, there will be no need to monitor every Tweet (or update) until a relevant one is posted. Instead, traffic influencers will have an intrinsic way of causing variations in road segments and will be less computationally extensive.
4.2.2. Traffic Prediction

Two tools will be used to perform the traffic prediction with information collected from historical data and social network analysis: Graph Convolutional and Long-Short Term Memory Neural Networks. These duos have been used in literature for traffic conditions, flow, or speed prediction [18,34,39]. However, none include a complete set of weather factors and their interaction with social media networks or POIs.

Graph Convolutional Neural Network (GCN)

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 7.

Figure 6. Example of COIN on Twitter with users interested in traffic within Los Angeles. Users are grouped in similarity clusters and the nodes are sized by influence (trendsetters).

Figure 7. Graph convolutional network architecture.
GCNs use forward propagation using an adjacency matrix to include graph information in the formula. Formula (1) 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 (2) is the adjusted formula for GCN:

\[
H^{l+1} = \sigma \left( W^l H^l + b^l \right) \\
H^{l+1} = \sigma \left( D^{-\frac{1}{2}} \tilde{A} D^{-\frac{1}{2}} W^l H^l + b^l \right)
\]

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.

Long-Short-Term Memory Neural Network (LSTM)

Long-Short Term Memory (LSTM) Neural Networks are 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 8, 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 (3), (4), and (5), respectively:

\[
i^{(t)} = \sigma \left( W^i \left[ h^{(t-1)}, x^t \right] + b^i \right) \\
f^{(t)} = \sigma \left( W^f \left[ h^{(t-1)}, x^t \right] + b^f \right) \\
o^{(t)} = \sigma \left( W^o \left[ h^{(t-1)}, x^t \right] + b^o \right)
\]

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 in Figure 8), 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.

The system architecture for this traffic prediction is shown in Figure 9, taking the model from [34] as a base but implementing POIs and traffic influencer analysis.
Figure 8. LSTM cell in RNN.

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 (3), (4), and (5), respectively:

\[
\begin{align*}
\text{i}(t) &= \sigma \left( W_{ij}(h(t-1), x_t) + b_j \right) \\
\text{f}(t) &= \sigma \left( W_{if}(h(t-1), x_t) + b_f \right) \\
\text{o}(t) &= \sigma \left( W_{io}(h(t-1), x_t) + b_o \right)
\end{align*}
\]

where \(\sigma\) is the activation function, \(W\) represents the weights, \(h(t-1)\) is the input from the previous computation, \(x_t\) is the input from the first layer (blue circle in Figure 8), 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.

The system architecture for this traffic prediction is shown in Figure 9, taking the model from [34] as a base but implementing POIs and traffic influencer analysis.

4.3. Illustrative Example Development and Validation

4.3.1. Algorithms Validation

The first step is a validation of the traffic prediction. Since there is limited data, the algorithm will first be validated by cross-validation. Cross-validation is a re-sampling technique commonly used for machine learning tools, called k-fold cross-validation, where \(k\) is the number of samples taken from the training dataset. This process will allow for a grid search for optimal parameters in the functions used, and it will determine if the further prediction of the validation set is underfitting, acceptable, or overfitting the training dataset.

Floating car data will be collected to validate traffic prediction under specific weather conditions. An example of the data collection table is shown in Table 5.

Table 5. Example of weather data (Los Angeles).

<table>
<thead>
<tr>
<th>Date</th>
<th>Wind</th>
<th>Visibility</th>
<th>Temperature</th>
<th>Dew Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-01-01T00:00:00</td>
<td>230</td>
<td>16,000</td>
<td>156</td>
<td>67</td>
</tr>
<tr>
<td>2019-01-01T00:53:00</td>
<td>170</td>
<td>16,093</td>
<td>156</td>
<td>56</td>
</tr>
<tr>
<td>2019-01-01T01:53:00</td>
<td>280</td>
<td>16,093</td>
<td>150</td>
<td>56</td>
</tr>
<tr>
<td>2019-01-01T02:53:00</td>
<td>30</td>
<td>16,093</td>
<td>150</td>
<td>−133</td>
</tr>
<tr>
<td>2019-01-01T03:53:00</td>
<td>110</td>
<td>16,093</td>
<td>139</td>
<td>−111</td>
</tr>
<tr>
<td>2019-01-01T04:53:00</td>
<td>110</td>
<td>16,093</td>
<td>122</td>
<td>−128</td>
</tr>
<tr>
<td>2019-01-01T05:53:00</td>
<td>110</td>
<td>16,093</td>
<td>122</td>
<td>−122</td>
</tr>
</tbody>
</table>

4.3.2. Case Studies and Validation of Results

For Case Study A, the methodology and framework will be applied to a megacity in the United States. This megacity will provide traffic and weather information and published vehicle routes from Amazon. This application will show traffic prediction accuracy under different conditions and why drivers yield further optimized routes.

Case Study B applies to a megacity resembling the Latin American urban infrastructure in the United States. This case study will set an example for using the framework in cities where public data is not widely available due to the city’s resource limitations. Furthermore, this work will show the importance of social media in the framework since it will be the main factor for traffic prediction. Latin American countries have a grid-like urban road network, as shown in Figure 10, which can be analogous to cities such as Orlando, Miami, or New York City, given their high population of Latino communities.
4.3. Illustrative Example Development and Validation

4.3.1. Algorithms Validation

The first step is a validation of the traffic prediction. Since there is limited data, the algorithm will first be validated by cross-validation. Cross-validation is a re-sampling technique commonly used for machine learning tools, called k-fold cross-validation, where k is the number of samples taken from the training dataset. This process will allow for a grid search for optimal parameters in the functions used, and it will determine if the further prediction of the validation set is underfitting, acceptable, or overfitting the training dataset.

Floating car data will be collected to validate traffic prediction under specific weather conditions. An example of the data collection table is shown in Table 5.

<table>
<thead>
<tr>
<th>Date</th>
<th>Wind Visibility</th>
<th>Temperature</th>
<th>Dew Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019-01-01T00:00:00</td>
<td>230</td>
<td>16,000</td>
<td>156</td>
</tr>
<tr>
<td>2019-01-01T00:53:00</td>
<td>170</td>
<td>16,093</td>
<td>156</td>
</tr>
<tr>
<td>2019-01-01T01:53:00</td>
<td>280</td>
<td>16,093</td>
<td>150</td>
</tr>
<tr>
<td>2019-01-01T02:53:00</td>
<td>30</td>
<td>16,093</td>
<td>150−133</td>
</tr>
<tr>
<td>2019-01-01T03:53:00</td>
<td>110</td>
<td>16,093</td>
<td>139−111</td>
</tr>
<tr>
<td>2019-01-01T04:53:00</td>
<td>110</td>
<td>16,093</td>
<td>122−128</td>
</tr>
<tr>
<td>2019-01-01T05:53:00</td>
<td>110</td>
<td>16,093</td>
<td>122−122</td>
</tr>
</tbody>
</table>

4.3.2. Case Studies and Validation of Results

For Case Study A, the methodology and framework will be applied to a megacity in the United States. This megacity will provide traffic and weather information and published vehicle routes from Amazon. This application will show traffic prediction accuracy under different conditions and why drivers yield further optimized routes.

Case Study B applies to a megacity resembling the Latin American urban infrastructure in the United States. This case study will set an example for using the framework in cities where public data is not widely available due to the city’s resource limitations. Furthermore, this work will show the importance of social media in the framework since it will be the main factor for traffic prediction. Latin American countries have a grid-like urban road network, as shown in Figure 10, which can be analogous to cities such as Orlando, Miami, or New York City, given their high population of Latino communities.

Results will be validated against vehicle routing and traffic datasets in the respective countries. The proposed framework will have to match these routes with 98% accuracy or yield further optimized ones. In addition, the framework will be validated by routing data from Amazon and another retail company. Amazon released a dataset for the routing problem, where the city and customer locations are known. The dataset has more than 4000 driver-determined routes, representing their expertise in the delivery areas and environmental factors that come into play.

The illustrative example for the traffic prediction methodology uses data from California’s Performance Measuring System (PeMS) District 7. The data is aggregated every 5 min for the weekdays of May and June of 2012. Data preprocessing has been performed following steps from [70] and further narrowed down to a subset of 25 roads. Preliminary experiments show that using an extended version of this neural network, called Bidirectional LSTM (BiLSTM), can help increase the accuracy of long-term traffic predictions where trends are cyclic. For example, Table 6 shows the mean absolute error (MAE) for experiments performed using traffic and road network data from 26 roads in California’s District 7 to predict traffic for the next 1 h, 6 h, 12 h, and 24 h.

Figure 10. The road network from cities in Latin America.
problem, where the city and customer locations are known. The dataset has more than 4000 driver-determined routes, representing their expertise in the delivery areas and environmental factors that come into play.

The illustrative example for the traffic prediction methodology uses data from California’s Performance Measuring System (PeMS) District 7. The data is aggregated every 5 min for the weekdays of May and June of 2012. Data preprocessing has been performed following steps from [70] and further narrowed down to a subset of 25 roads.

Preliminary experiments show that using an extended version of this neural network, called Bidirectional LSTM (BiLSTM), can help increase the accuracy of long-term traffic predictions where trends are cyclic. For example, Table 6 shows the mean absolute error (MAE) for experiments performed using traffic and road network data from 26 roads in California’s District 7 to predict traffic for the next 1 h, 6 h, 12 h, and 24 h.

Table 6. California’s District 7 traffic prediction experiments using [70] as a benchmark.

<table>
<thead>
<tr>
<th>Time</th>
<th>GCN-LSTM MAE [70]</th>
<th>GCN-BiLSTM MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 h</td>
<td>0.13</td>
<td>0.161</td>
</tr>
<tr>
<td>6 h</td>
<td>397,956.564</td>
<td>0.161</td>
</tr>
<tr>
<td>12 h</td>
<td>84,232.415</td>
<td>5,451,617.781</td>
</tr>
<tr>
<td>24 h</td>
<td>5.48 \times 10^{23}</td>
<td>1.5</td>
</tr>
</tbody>
</table>

5. Discussion

As evidenced in the literature review, traffic flow refers to the number of vehicles passing through a particular road segment during a specified period. Traffic flow is an essential factor influencing the travel time of vehicles. It has been seen that in general, a high traffic flow leads to more congestion and slower travel times, while a low traffic flow results in faster travel times. For vehicle routing, it is crucial to consider traffic flow because it directly affects the travel time of vehicles. If a routing algorithm only considers the shortest distance between two points without taking traffic flow into account, it can lead to inefficient routing decisions. For example, a route that appears to be the shortest distance on the map may have a higher traffic flow, resulting in longer travel times.

Predicting traffic flow may not be the most effective approach to determine the fastest route for vehicle routing but it does help to identify the prediction of the time it would take a vehicle to go from one place to another. For example, in practical cases, we have discussed how traffic speed or travel time prediction can be achieved by using machine learning algorithms that take into account various factors such as historical traffic patterns, time of day, weather conditions, and the characteristics of the road, the road network, and others. These algorithms can provide real-time information on traffic conditions, allowing vehicles to adjust their routes to avoid congestion and reduce travel times. In short, understanding traffic flow is essential for vehicle routing as it directly affects travel times. However, predicting traffic speed or travel time may be a more effective approach than predicting traffic flow when determining the fastest route for vehicle routing.

6. Conclusions and Future Research

The conclusion of this paper highlights the significance of the proposed methodology in traffic prediction and dynamic vehicle routing problems. The combination of deep learning and social media analytics is expected to bring a new perspective to real-time dynamic vehicle routing problems and improve decision making. The case study results show that applying deep learning techniques can lead to more accurate and efficient predictions. At the same time, integrating social media analytics, point of interest analysis, and network analysis allows for the adaptation of the solution to various scenarios. Some limitations of this proposed methodology would be computational time, since GCNs can be very time demanding as the dataset grows, and there is a lack of data availability for underdeveloped countries (for example, some cities in Latin America). However, these
two issues can be resolved in future studies by implementing an attention mechanism and using data from cities that portray a similar lifestyle to Latin America, respectively.

Future research should address the challenges of urban last-mile delivery problems in developing countries with limited Intelligent Transportation Systems data. This study is a crucial first step in developing a framework to support these countries in improving their logistics systems. Further research should aim to identify the most influential variables affecting traffic prediction and the impact of social media and other external factors on traffic patterns.

Moreover, this framework provides a platform for further advancements in traffic prediction and vehicle routing. It opens the door for new and innovative methods for predicting traffic patterns and optimizing delivery routes. For example, the co-occurrence word graph used in this study can be expanded to include more terms and variables, providing a more comprehensive and nuanced view of the relationships between key concepts in the field. Additionally, the proposed methodology can be modified to account for other factors impacting traffic prediction, such as weather conditions, road construction, etc.

In conclusion, the proposed framework for urban last-mile delivery traffic forecasting using social media analytics and deep learning represents a significant contribution to traffic prediction and dynamic vehicle routing problem research. The literature review and case studies indicate the potential of this methodology to bring improved accuracy and efficiency to traffic prediction and to provide new insights into the complex problem of last-mile delivery in urban areas. However, further research is needed to fully realize this framework’s potential and continue to advance our understanding of traffic prediction and vehicle routing.

Author Contributions: Conceptualization, V.L.-F.; methodology, V.L.-F. and E.G.-F.; theoretical analysis, V.L.-F. and L.R.; validation, E.G.-F, L.R. and G.L.; writing—original draft preparation, V.L.-F. and E.G.-F.; writing—review and editing, A.T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References


16. Ma, Y.; Zhang, Z.; Ihler, A. Multi-Lane Short-Term Traffic Forecasting With Convolutional LSTM Network. IEEE Access 2020, 8, 34629–34643. [CrossRef]


25. Huang, M.-L. Intersection traffic flow forecasting based on N-GSVR with a new hybrid evolutionary algorithm. Neurocomputing 2015, 147, 343–349. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.