International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Predictive Modelling of Traffic Flow with Deep Learning Techniques

[1]M. Srujan Kumar, [2] M. Sruthi, [3] M. Suhaas
[4]M.Suryam, [5]C. Vasanth Kalyan,
[6] Prof R. A. Manikandan
[12345]Student, [6] Assistant Professor

[1] 2111cs020569@mallareddyuniversity.ac.in, [2] 2111cs020570@mallareddyuniversity.ac.in,
[3]2111cs020574@mallareddyuniversity.ac.in, [4] 2111cs020580@mallareddyuniversity.ac.in,
[5]2111cs020527@mallareddyuniversity.ac.in

Department of Artificial Intelligence & Machine Learning
Malla Reddy University, Hyderabad.

Abstract:

This project delves into predictive modeling for traffic flow using deep learning techniques, focusing on the Metro Interstate dataset. Traffic Flow Prediction (TFP) is crucial for Intelligent Transport Systems (ITS), optimizing vehicle movement, reducing congestion, and improving route efficiency. Leveraging advancements in Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Big Data, our study explores various techniques and models for TFP. We highlight DL models' advantages over traditional ML methods, propelled by the wealth of real-time traffic data fostered by smart cities, presenting opportunities to craft robust predictive models. The core of our project revolves around developing a multi-step Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) architecture. Our model forecasts traffic volume between Minneapolis and St. Paul, Minnesota, predicting volume two hours into the future based on a six-hour historical window. We explore DL algorithms' efficacy, including LSTM and Gated Recurrent Unit (GRU), in mitigating challenges like the vanishing gradient problem common in RNNs. Our analysis compares various NN models, emphasizing the importance of data availability for training and fine-tuning ML/DL models, with the Metro Interstate dataset serving as a crucial asset for comprehensive traffic flow analysis and model development.

1. INTRODUCTION

Traffic flow forecasting (TFP) mainly refers to estimating the volume and speed of traffic using the method of controlling the movement of vehicles, reducing traffic congestion and creating favorable conditions (less time or energy consumption). With recent advances in artificial intelligence, machine learning (ML), deep learning (DL), and big data, traffic prediction research has also expanded. TFP is an important part of Intelligent Transportation Systems (ITS) and can assist ITS in traffic forecasting. Traffic regulations are very difficult in big cities. Many countries have adopted efficient transportation systems to reduce transportation costs. In this study, the applications of artificial neural network (ANN), ML, DL and other technologies and models in TFV are examined. Finally, we will create our own prediction model using deep learning, train and test it, verify its accuracy, and compare the accuracy of our model with other models. Compared with traditional machine learning methods, deep learning models have the advantage of primarily simplifying data and being more accurate than other machine learning methods. Therefore, data driven traffic prediction has attracted much attention in TFP recently due to the large amount of data and deep learning based on previous data. In addition, with the development of smart cities in recent years, the available information has also increased. This innovation can benefit transportation in the field of smart transportation, reducing travel times, increasing efficiency and reducing the environmental impact of vehicles.

Machine learning and deep learning are rapidly growing fields for traffic prediction. Traffic signs, traffic jams, weather and traffic control are the main causes of traffic congestion. As the size of real- data traffic increases, big data should be used to improve data transmission. This fact brings us between Minneapolis and St. Paul from special places in Minnesota. Our goal is to build a multistep recurrent neural network (RNN) with a short term temporal (LSTM) model to predict traffic in the next 2 hours, given 6hour window transitions.

2. LITERATURE REVIEW

The literature review on Traffic Flow Forecasting (TFP) and Intelligent Transportation Systems (ITS) includes extensive research of secondary sources such as articles, books, and research papers. Most studies on TFP focus on short term traffic forecasts, often using historical data. Taking into account instantaneous estimates of traffic flow, the differences between TFP and regional management in urban transportation are revealed. Various artificial intelligence (AI) technologies, especially machine learning (ML) and deep learning (DL) models, can be used for accurate traffic prediction, helping reduce accidents and improving traffic control.
Recent advances in TFP include the use of graph neural networks (GNN) for traffic prediction, the use of rich mobile phone data, and deep learning algorithms. Integration of machine learning techniques with online and offline traffic data has made progress in traffic prediction. Studies by researchers such as Karthika, Chandel, Singh, Jiang, and Khiewwan exploring different ML/DL models, regression techniques, data mining, and hyperparameter transformation techniques demonstrate continuous efforts to improve the accuracy and effectiveness of TFP models to increase efficiency.

3. METHODOLOGY

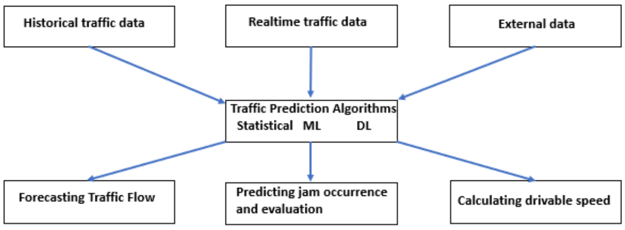

Fig. 3 Approaches to Traffic Flow

3.1 Data Collection

The Metro Interstate dataset provides hourly traffic data for the westbound lanes of U.S. Interstate 94 (I94) from 2012 to 2018. Weather and holidays. Prerequisite include cleaning, handling missing values, statistical modeling, and entering categorical variables to validate the model.

3.2. Model Architecture

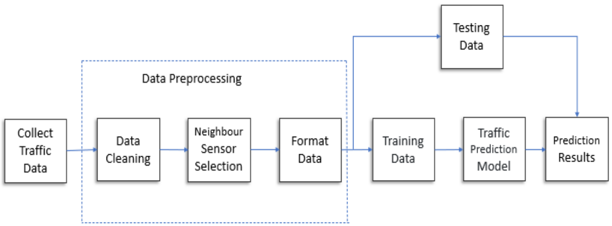

Fig. 3.2 Architecture

3.3. Feature Engineering

Feature engineering for traffic prediction involves the use of deep learning methods such as convolutional neural networks (CNN) and recurrent neural networks (RNN) such as short term (LSTM) and Gated Recurrent Unit to predict traffic congestion based on 2D data matrices. (GRU) models can solve the problem of gradient disappearance and improve the accuracy of traffic prediction. It is also recommended to use other neural network models such as graph neural networks and hybrid techniques to improve the prediction.

Larger data can help achieve higher accuracy in ML/DL algorithms for traffic prediction.

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 48204 entries, 2012-10-02 09:00:00 to 2018-09-30 23:00:00
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   holiday         48204 non-null  object
 1   temp            48204 non-null  float64
 2   rain_1h         48204 non-null  float64
 3   snow_1h         48204 non-null  float64
 4   clouds_all      48204 non-null  int64
 5   weather_main    48204 non-null  object
 6   traffic_volume  48204 non-null  int64
dtypes: float64(3), int64(2), object(2)
memory usage: 3.9+ MB
```

Fig 3.3 Data Features and Types Overview

Data Preprocessing:

Preparing data for the machine learning model is what we call preliminary data and is an important first step in building the model. In real world machine learning projects, we often don't start with clean data. It's like a mess that needs to be sorted out before it can be put together. Every time we work with data, it needs to be cleaned and created correctly. This is where data preprocessing comes into play; cleans and organizes data so it's ready for the machine learning magic to happen.

Model Building:

Deep learning, especially neural networks such as CNNs and RNNs, perform well in predicting traffic with over 90% accuracy. Different methods such as LSTM and GRU address issues such as incompleteness. To achieve the best accuracy in traffic prediction, it is recommended to use different neural network models and hybrid methods that leverage larger data for training and optimization.

Training the Models

LSTM:
Short term traffic forecasting is essential for smart transportation. Future studies will include additional spatial and temporal data to improve the accuracy of larger routes. Increasing the number of LSTM layers can improve learning, but also increases the risk of overfitting.
CNN:
CNN uses a filter layer to remove features. When backpropagation controls the weight variation, CNN provides parallel filters to extract different features, thus improving the multithreading model.
CNN VS LSTM:
CNN combines signal processing with deep learning through reduced but controlled processing. Although LSTM is parameter intensive, it is good at analyzing long sequences without the need for network expansion.
LSTM RNN:
LSTM solves RNN's loss problem by allowing cells to forget information, making it ideal for situations such as traffic prediction. LSTM RNN outperforms nonparametric models in this area.

3.4 Test Results and Analysis

Metro Interstate Traffic Dataset, Minneapolis and St. Paul, Minnesota. Hourly and weekend weather features are included to understand the impact on traffic. Our main goal is to use an LSTM model to create a multi-level RNN to predict traffic in the next 2 hours based on the previous 6hour window.
This can be illustrated in (Figure) where the input represents the data points issued in 6 hours and the label represents the demand in 2 hours.
Check in hours. It measures real time series data, but this does not yield useful results (average MAE of 300 seconds).
We then decided to effectively preprocess the training and data into a time series indexed at 1hour intervals, and the results were better (less than 200 seconds at this MAE). . Follow records that are accurate records.
There are gaps in the data as shown in the image below.

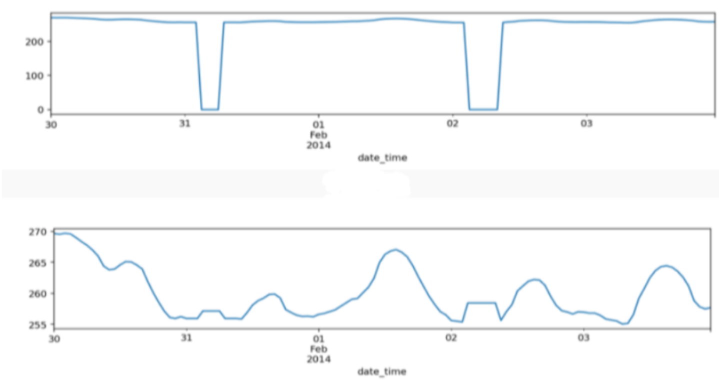

Fig 3.4 Effect of Data Preprocessing on Validation

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Results

3.5 Prediction Evaluation

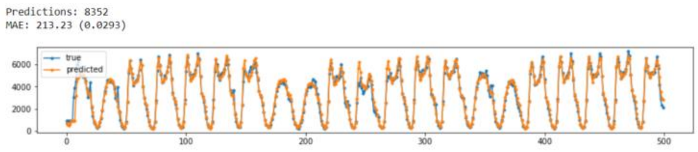

Fig 3.5 Evaluation of LSTM Model Variants.

4. CONCLUSION

Today, cyber attacks pose a serious threat that leads to service disruptions and leakage or misuse of important user data. This article addresses the fundamental problem of unauthorized content in wireless LAN networks that can disrupt service and access the user's important information without the need for special knowledge. Rapid detection and mitigation of these threats is important to prevent harm resulting from their use. User authentication is initiated by malicious actors. These malicious details allow attackers to impersonate legitimate users, making it easier to intercept communications and data., K Nearest Neighbors (KNN) and Multilayer Perceptron (MLP) – because they are very effective in detecting irregular points. However, it is accepted that requirements such as scalability and seamless integration into the existing network infrastructure are important for the effectiveness of protection blocks against evolving cyber threats.

4.2. References

1. Conducting research on traffic prediction using deep learning machine. International Journal of Innovative Technology and Exploration Engineering, 2019.

2. Apurv Chandel, S.S., Badavath Uday Kiran, Prabhas Prasad, Nidhi Lal, Accurate prediction of traffic in urban areas using machine learning linear regression model.

3. Ahmad, United Kingdom Metro interstate traffic time series forecasting using Recurrent neural networks(RNN).Availableat: https://medium.com/@umaimakhurshidahmad/metro-interstate-traffic-volume-time-series-forecasting-using-recurrent-for-neural-networks-rnns-a73732276d1a.

4.Garlan,E.;Availableat: https://www.kaggle.com/code/ramyahr/metro-interstate-trafficvolume/notebook.

5. Rohit Singh, Mr. Sibi Amaran, Dr. K Sree Kumar, Traffic analysis based on weather variables. April 2021.

6. Jiang, R. et al., DL-Traff: A survey and deep learning model for urban traffic prediction, Proceedings of the 30th ACM International Conference on Information and Knowledge Management. 2021. 4515-4525 Ib.