Exploring Data Science and Machine Learning (Smart City Traffic Pattern)

Under

UpSkill CAMPUS

Submitted by-
Shivam

Date - 12-07-2023

Duration - (1 June to 15 July)

Table of Contents:

Smart City Traffic Pattern Summary:

I. Objectives:

1. Improve traffic flow and reduce the number of people in the city.
2. Improve transportation and reduce travel time.
3. Improve the use of existing systems and reduce the need for innovation.
4. Increase the safety of pedestrians, cyclists and drivers.
5. Promote sustainable transportation and reduce carbon emissions.

II. Methodology:

1. Leveraging advanced technologies such as the Internet of Things (IoT), machine learning, and data analytics.
2. Use high-speed traffic management to monitor and analyze traffic patterns in real time. The collects data from a variety of sources, including sensors, cameras, and vehicle connectivity.
3. Use predictive models and classification techniques to predict traffic patterns and predict potential problems.
4. Adjust signal timings, implement traffic control strategies and improve traffic flow.
5. Collaborate with other smart cities such as smart transportation and power management to create an integrated system.

III. Key findings:

1. A new model identifies and analyzes IoT smart city use cases in five modes of transportation
2. Smart city planning uses a classification system that focuses on data collection and data mining for traffic forecasting
3. Machine Learning Algorithms are developed to predict traffic patterns in smart cities used successfully
4. Intelligent traffic management has the potential to increase the capacity of urban roads without adding new roads
5. Intelligent railways and traffic management have changed urban traffic management and emergency response
6. Integrated systems and management Machines with nerve impulses are used urban traffic forecasting and smart city management to

IV. Conclusion

Smart city traffic models aim to improve traffic flow, reduce accidents and improve traffic in the city. By leveraging advanced technologies and a data-driven approach, the city will improve mobility, increase security and create a better transport system. Continuous research and development in this field continues to introduce new solutions for smart city traffic management.

1. Introduction

In the age of urbanization and technological progress, the concept of smart city has emerged to increase the quality of life of the inhabitants. An important aspect of smart cities is the use of traffic management systems. These systems use advanced technology and data-driven optimization to improve vehicle models and improve transportation as a whole.

Smart city train models are predicted and controlled using a variety of technologies, including machine learning and intelligent traffic management. This technology monitors traffic flow, identifies patterns and adjusts signal timing to reduce congestion and improve the overall experience

City traffic management is crucial to creating a smart, sustainable, efficient city. Using technology and a data-driven approach, smart cities can improve traffic patterns, improve mobility and improve overall quality of life for residents.

The goal of smart city traffic management is to create a better and more efficient transportation system without the need for major expansion. The smart city will increase traffic flow, increase security and optimize existing infrastructure by leveraging connected vehicle technology and leveraging a consistent set of hardware, software and cloud solutions.

2. Methodology

a) DataSet:

Smart city traffic patterns dataset is collected from the Kaggle.The file has 48, 120 lines and just four column like DateTime, Vehicle, Junction and ID.

	DateTime	Junction	Vehicles	ID
0	2015-11-01 00:00:00	1	15	20151101001
1	2015-11-01 01:00:00	1	13	20151101011
2	2015-11-01 02:00:00	1	10	20151101021
3	2015-11-01 03:00:00	1	7	20151101031
4	2015-11-01 04:00:00	1	9	20151101041

b) Data Preprocessing :

It involves cleaning and manipulating the collected data to make it good and efficient with machine learning algorithms. Below is a description of the data before the steps under the smart city traffic pattern analysis:

i. Data Cleaning: This step involves removing irrelevant or duplicate content from the data. It also includes handling of missing values, which can be done by methods such as assignment or deletion depending on the missing data.

ii. Data integration: Data on traffic patterns in smart cities can be collected from a variety of sources and sensors. Data integration involves combining these disparate data into a unified framework for analysis. This step ensures that all relevant information is included and can be used together to gain insight.
iii. Data Conversion: Data conversion converts the collected data into a format suitable for analysis. This may include converting categorical variables to numerical representations, scaling numerical variables to multiples, or using numerical transformations to normalize data distributions.
iv. Feature Engineering: Feature engineering is the process of generating new features from existing data that can increase the predictive power of machine learning models. In the context of smart city traffic models, this may involve extracting physical parameters such as time of day, day of the week or season from time data. It may also include providing spatial features based on available data.

C) Machine Learning Algorithm:

i. SVMs: The SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.We evaluate the SVM model with linear kernels.

ii. Decision Tree: The goal of Decision Tree is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piecewise constant approximation.Decision tree creates biased trees if some classes dominate. Therefore we tried to balance the dataset prior to fitting with the decision tree but Decision trees can be unstable because small variations in the data might result in a completely different tree being generated.

iii. Random Forest: Random Forest works in two-phase first is to create the random forest by combining N decision tree, and second is to make predictions for each tree created in the first phase.Select random K data points from the training set.Build the decision trees associated with the selected data points (Subsets).Choose the number N for decision trees that you want to build

iv. Logistic Regression : Logistic regression predicts the output of a categorical dependent variable. Therefore the outcomemust be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but insteadof giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.

v. RNN : A Recurrent Neural Network (RNN) is a type of artificial neural network that is specifically designed to process sequential data or time-series data. Unlike traditional feed-forward neural networks, RNNs have the ability to retain information from previous inputs, allowing them to exhibit temporal dynamic behavior. This makes RNNs well-suited for tasks that involve processing sequences of inputs, such as speech recognition, natural language processing, and smart city traffic pattern analysis.

vi. XGBoost: XGBoost, short for Extreme Gradient Boosting, is a powerful and widely used machine learning algorithm popular in machine learning and Kaggle competitions on datasets or databases. It is an implementation of Gradient Boosted Decision Trees designed for speed, efficiency and scalability.

3. Result and Analysis:

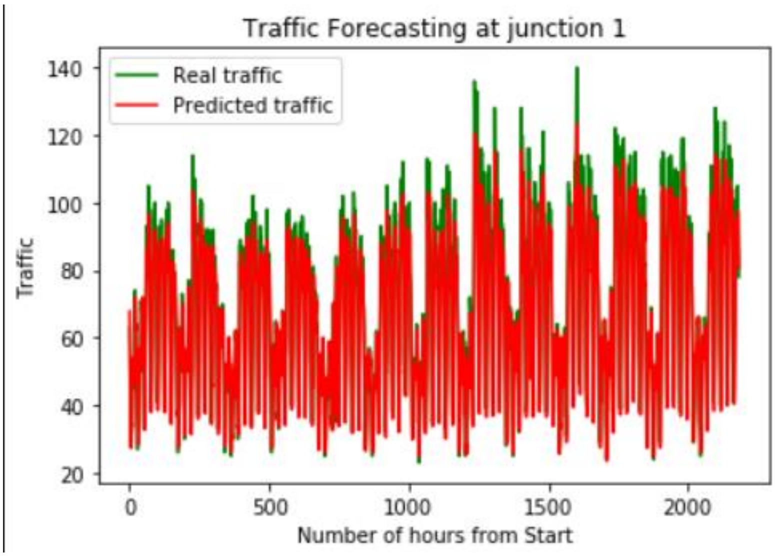
In this research we have applied 5 algorithm SVM, Decision Tree, Random Forest, XGBoost, RNN. The Highest Accuracy we can achieve with decision tree model. The below table shows respective results with there comparison.

Algorithm	Accuracy
Random Forest	19.00
SVM	8.30
XGBoost	32.00
Decision Tree	100.00

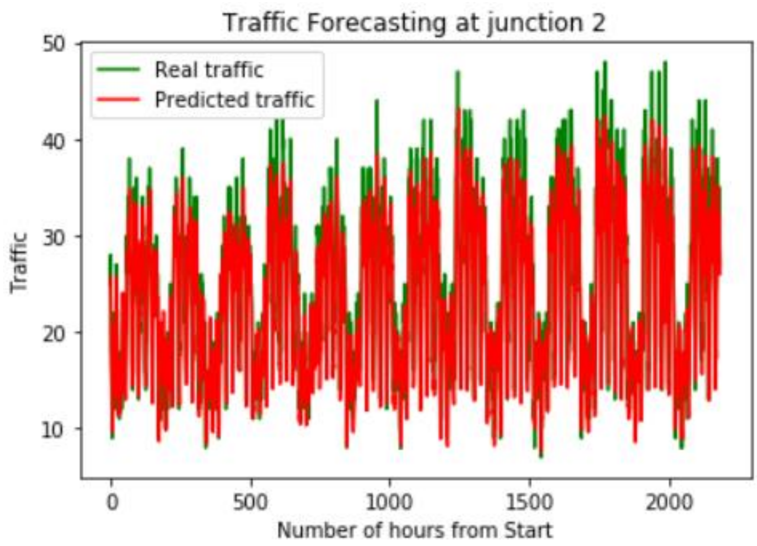
Table 1. Result comparison

Traffic Forecast of all Junctions:

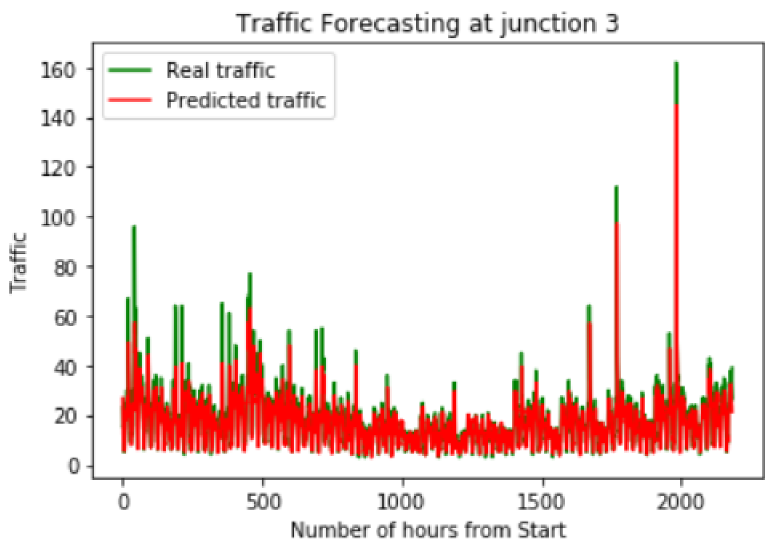
i. Junction 1:



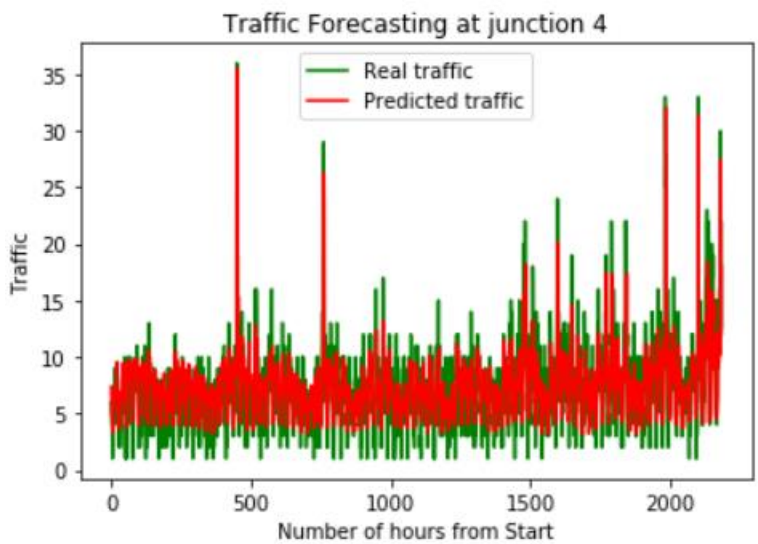
ii. Junction 2:



iii. Junction 3:



iv. Junction 4:



4. Discussion:

I. Challenges and limitations:

  i. Limited data availability: One of the challenges of analyzing traffic patterns in smart cities during operation may be the availability of good data and Good information. Obtaining real-time, accurate data from multiple sources can be difficult, especially when there are limitations on data collection or access to relevant information.

  ii. Data quality and consistency: Maintaining the quality and consistency of data collection can be another challenge. The information may contain errors, omissions or inconsistencies that should be addressed at the earliest opportunity. Resolving such issues can be time consuming and affect the accuracy of the analysis.

II. Pros and Cons of Approach:

  I. Pros:

    i. Using machine learning algorithms such as SVM or RNN can provide accurate predictions and insights into traffic patterns in smart cities.

    ii. Advanced procedures such as data cleaning and optimization can improve the quality and accuracy of data used for analysis. Theuses advanced algorithms and models to capture complex patterns and relationships in traffic data.

  II. Weaknesses:

    i. This method will rely heavily on the availability and quality of data. Insufficient or inaccurate data may limit the validity of the analysis.

    ii. Algorithm options may be limited in processing certain types of data or capturing certain patterns in traffic data.

    iii. Roads may not include other factors that also affect traffic patterns, such as city planning or policy interventions.

III. **Future Work:**

  i. Improve data collection: Invest in improving data collection processes and procedures to ensure data quality and the ability to analyze smart city traffic patterns.

  ii. Inclusion of External Factors: Consider including other factors in the analysis, such as city planning or policy interventions, to better understand traffic patterns and drivers.

  iii. Research Methods: Explore the use of methods such as RNN to combine predictions of various models to increase the accuracy and power of model predictions.

  iv. Real-Time Analytics: Build real-time analytics capabilities to help tailor and adapt traffic management policies to updated forecasts.

  v. Collaboration: Collaborate with local governments, transport authorities and urban planners to gain relevant knowledge and expertise and facilitate dealing with real needs and problems.

5. Conclusion:

Consequently, analyzing traffic patterns in smart cities using machine learning techniques has the potential to improve transportation and improve the overall quality of life in cities. Using data collected from IoT sensors and other sources, machine learning algorithms can predict traffic patterns, improve public transport systems, and develop traffic efficiency control strategies.

However, there are still some challenges and limitations that need to be addressed. These include the availability and quality of data, the complexity of traffic patterns, and the need for data collection. In addition, the analysis should consider other factors such as urban planning and policy interventions to better understand traffic patterns.

The strength of this method lies in its ability to capture complex patterns and relationships in data using sophisticated algorithms and advanced methods to improve the quality of the data.

To improve future research and use of traffic measurement models in smart cities, it is recommended to strengthen the data collection process, participate in other aspects of analysis, integrate search, develop real-time inspection capabilities, and establish partnerships with stakeholders. .

By addressing these recommendations, future efforts to analyze traffic patterns in smart cities can lead to more predictive, effective traffic management strategies and ultimately the development of smart and liveable cities.