WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Road Traffic Anomaly Detection Using Deep Learning

Nehal Pawar

AISSMS IOIT, Pune

A B S T R A C T

The advancement of 5G has made it possible for autonomous vehicles (AVs) to have total control over every aspect of operation. The AV operates on its own and gathers trip data based on various smart sensors and detectors in order to be able to run on its own. Nevertheless, various factors that impair the soothsaying delicacy have an impact on the data that has been gathered. A computational data wisdom strategy (CDS) is suggested for managing vast amounts of commercial data in various formats. The computational data wisdom system was developed to identify data anomalies that impair company effectiveness. Advanced artificial intelligence techniques, such as deep learning, combined with data wisdom enable the advanced degree of data anomalies discovery that reduces business traffic and vehicle queuing. The primary contribution of the CDS strategy is exemplified in the early identification of the causes of data anomalies in order to prevent long-term business locks. Additionally, CDS displayed a promoting outcome in vibrant road business scripts. A larger family of machine learning techniques built on artificial neural networks and representation learning includes deep learning.

Keywords: Deep Learning, CDS-Computational Data Science

Introduction

The detection and classification of road traffic anomalies is main focus of this research. In regard to the identification of elements that characterize the road conditions, besides other surface anomalies, environment and health hazards have been identified. Road anomaly detection and road condition survey systems based on collaborative mobile sensing typically detect and automatically classify road anomalies by applying data-mining approaches on data collected by smartphones. The objective was to evaluate a road anomaly detection system based on a machine learning. A deep learning model is suggested here to manage the dataset input. The dataset was analyzed based on a data science life cycle concept that comprises the data collection, data preparation, data exploration, modeling and model evaluation and deployment. A goal-oriented agent system fulfills designed functions, such as detection of data anomalies, identification of their sources and the localization of the traffic data anomalies based on statistical analysis. The proposed computational data science combines artificial intelligence and data were evaluated and assigned a score based on the quality life cycle

Motivation

The road traffic anomalies are the major problems in the big cities. The management of the larger data sets is quiet harder in real world deployment. To improve the larger data inputs and classify the anomalies into the different categories this study will be helpful. The proper categorization of anomalies will help to reduce the road accidents and traffic. Safety percentage while driving increases due to this sorting.

1.1 Urban traffic accident risk prediction for knowledge-based mobile multimedia service

The statistical methods have limitations in that, they are based on assumptions about data distribution and function type. Therefore, this study suggests an accident prediction model using deep learning. This newly suggested risk prediction model is for predicting risk by reflecting static features of the road, such its length and the speed limit on it, and dynamic features of the road, such as traffic volume when driving on it, and the altitude and azimuth

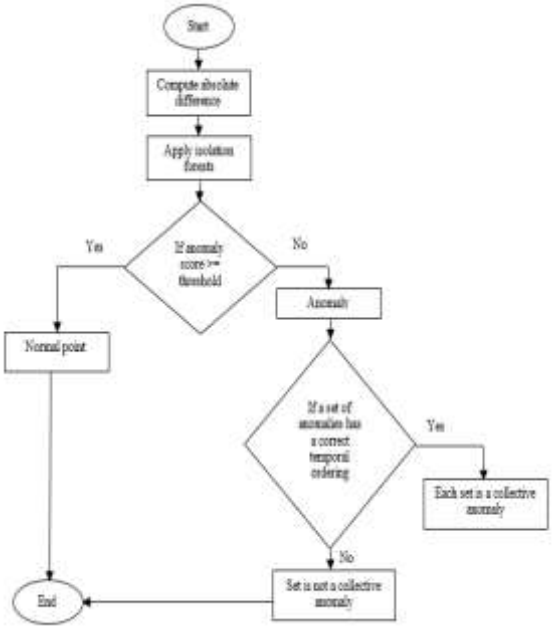
1.2 Radio Resource Management for D2D-based V2V Communication

Vehicle-to-vehicle (V2V) communications have attracted great interest due to the potential to improve traffic safety, reduce energy consumption, and enable new services related to intelligent transportation systems. Usually, these types of applications have a strongly localized nature, requiring cooperation between vehicles in close proximity.

1.3 Machine learning based optimization for vehicle-to-infrastructure communications

The study of wireless communications in vehicle-to-infrastructure communications is done in this paper. Multiple vehicles exchange information via a common roadside infrastructure. Investigation of methods to incorporate machine learning into the optimized system. Importance of machine learning can be adopted in various models to realize the optimized system performance is discussed.

2.1 Architecture



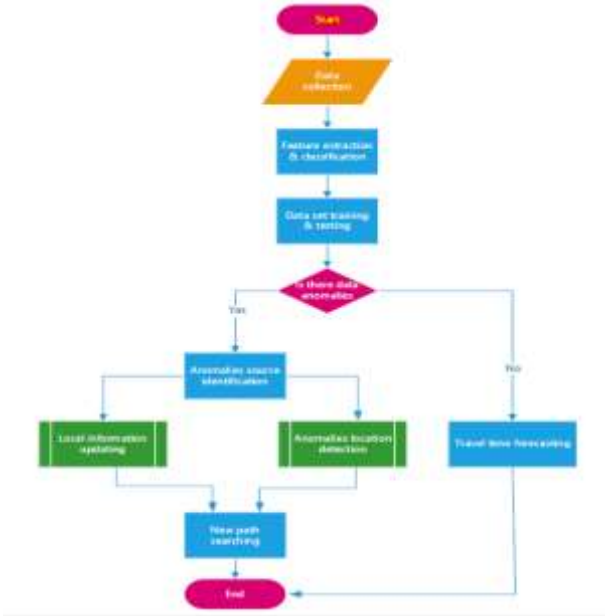


Fig (a) Default Architecture for anomaly Detection ***Fig (b) Proposed Architecture for Anomaly Detection***

3.1 Methodology

3.1.1 DL-Based Detection of Positioning Data Anomalies

Each device on an AV was subjected to a threshold of sensitivity analysis, which involved field tests of the GNSS-based positioning terminal (GBPT) carried out under real conditions to identify positioning errors. The field tests were specialized tests with an embedded reference trajectory measurement system for delivering the ground truth. This method had the advantages of expanding the number of field tests that could be executed under real operational conditions

The process of anomaly detection faces many challenges. It is needed for discovering a pattern that do not conform to normal expected behaviour in the data. The first step in the anomaly detection process is to define a normal traffic on a section of urban road and then to flag as anomalies any observations that are not fit with this normal pattern. The main challenge for anomaly detection in urban road traffic is to find these patterns. In comparison to traditional schemes, the CDS approaches can be applied to all types of normally trafficked roads. The isolation forest algorithm will work efficiently for this anomaly classification.

Many algorithms are proposed to detect the data anomalies in traffic road, their activities are directed toward influencing the data anomalies on road traffic which are expressed in traffic congestion. On other hand, the activities of CDS algorithm focused on data behaviour which is influence by various factors. The cognitive data differentiates between the influence of internal factors, such as delay in vehicle –to-vehicle communication that are caused due to variation of requirement for QoS, and extern factors that influence the data such as cyber attackers, geographical factors, radio channel interference.

3.1.2 Data Anomaly Detection Based on CDS

The travel data and AV positioning data were analysed with various techniques. The most critical issue in collecting the data was finding the right input data set. The big data input went through several preparatory steps:

- **Recognition data**

AVs collects various types of data from their installed devices. Classification tools are needed to identify the types of data sets. The data on vehicle positioning was collected from Ublox devices and smartphones.

• Structured vs. unstructured data

In general, AVs collect structured data. Vehicle positioning data is often represented by a matrix, whose columns represent distinct properties of these items in it. For instance, the set of positioning data input for the study contained one column for longitude and other columns for altitude. Some of the devices installed in Avs collect unstructured data, such as GNSS data, so our first step is to build a matrix to structure them.

• Cleaning and formatting

An important step is cleaning and formatting the data. Travel data are collected by different devices and are influenced by human and environmental factors. *Data cleaning* is the process of modifying the data to ensure that the datasets are free of irrelevancies and incorrect or incomplete travel observations. The best computational data formats have several useful properties; for instance, they are easy for computers to parse, easy for people to read, and widely usable by other tools and systems. The travel data input for the study came from various smart devices in all kinds of formats. This dataset was cleaned to improve the efficiency of the data analysis and the quality. This dataset was cleaned to improve the efficiency of the data analysis and the quality of the results.

4. ADVANTAGES AND DISADVANTAGES

4.1 Advantages

1. This project can classify the anomalies efficiently and detect them properly.
2. This will help to work on larger and real time data set .
3. It also solves problems like road reflection and multipath effects due to the environment.

4.2 Disadvantages

1. With certain benefits, there are also drawbacks, such as it will not work effectively if data is abnormal.
2. We cannot claim the assurance for totally safe and error free traffic detection

5. CONCLUSION

AV communications can be affected by various problems which we have discussed earlier. Thus anomalies can increase road accidents and congestion. AVs can experience bad propagation conditions for extended periods of time. The simulation results demonstrate that the forecasting system was improved by the use of the DL network. The original raw data was preprocessed. Missing information was compensated for with a data cleaning process that involved estimating the data from the moment before the gap. Missing information was compensated for with a data cleaning process that involved estimating the data from the moment before the gap. This reduced or eliminated unwanted features attributed to noise in the original data. The processed dataset was divided into training and testing subsets to carry out supervised learning. A system is proposed to detect anomalies in traffic flow that lead to congestion. The DL concept, based on statistical measurements, makes possible the early detection of traffic congestion and traffic accidents. Thus, the proposed system may have a direct and significant positive impact on driver's health and safety. The forecasting system is also improved by the use of the DL network