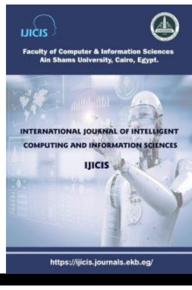
IJICIS, Vol.23, No.3, 114-140
DOI: 10.21608/ijicis.2023.192999.1254

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



DRIVER BEHAVIOR DETECTION IN TIME SERIES DECADE REVIEW

Rabab Gamal*
Computer Systems Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt
rabab.gamal@cis.asu.edu.eg

Mirvat Al-Qutt
Computer Systems Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt
mmalqutt@cis.asu.edu.eg

Said Ghoniemy
Computer Systems Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt
ghoniemy1@cis.asu.edu.eg

Received 2023-02-09; Revised 2023-02-09; Accepted 2023-07-12

Abstract: *Driver's behavior is expressed by the intentional and unintentional actions the driver performs while driving a motor vehicle. This behavior could be influenced by several factors such as fatigue, drowsiness, vehicle surroundings, and distraction state. Driver's behavior could be normal, risky or aggressive. Risky and aggressive behaviors, such as harsh braking and rapid acceleration, can lead to traffic accidents. Monitoring, analyzing and improving driver's behavior can reduce traffic collisions and enhance road safety. Different approaches have been followed for the detection and identification of driver's behavior. Rule-based Machine learning (ML) and deep learning (DL) approaches have succeeded to mine dynamical characteristics of time series. However, they have some challenges that make them unsuitable for many classification tasks including the selection of efficient architectures and corresponding hyper-parameters, as well as slow training and limited labeled data. Fusion and attention mechanisms through hybrid approaches were found to be more suitable for time series sensor data analysis. Transfer learning addresses useful approaches for making use of learning applied to other applications. Maneuver detection represents a serious characteristic of driver's behavior identification. A recent approach for extracting maneuvers from high-frequency telematics data is through time series motifs detection algorithms. Motifs extraction is preferred over ML and DL approaches as it does not requie labels, which is extremely time-consuming to collect. This work focuses on the latest techniques for classifying driver's behavior in time series data, and summarizes the pros and cons of the different categories.*

Keywords: *Driver Behavior Detection, Time-Series Data Analysis, Motifs, Machine Learning*

1. Introduction

Building Insights about vehicles driving behavior helps automotive manufacturers optimize their designs, simplify maintenance, and improve safety. The 2022 status report of the World Health Organization (WHO) [1] on road safety states that the lives of more than one million people are cut short yearly because

***Corresponding Author**: Rabab Gamal

Computer Systems Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt

Email address: rabab.gamal@cis.asu.edu.eg

of road traffic crashes. The traffic accidents in Egypt from 2015 to 2019 are presented in Table 1 according to the Central Agency for Public Mobilization and Statistics (CAPMAS) [2].

This excessive number of deceased and injured persons pushes the society to pay more attention for road safety which is ensured through a holistic view of the road system. Driver behavior is one of the most important players for road safety. Attitude, emotions, exhaustion, drowsiness, driving conditions, etc., are just a few examples of the many variables that can influence or change a driver's behavior.

Table 1 Number of car accidents, their results, and rates in the first half from 2015 to 2019

Year	***Number of accidents***	***Accident results***			***Accident Rates***		
		Deceased	***Injured***	***Damaged vehicles***	***Accident/Month***	***Deceased/Month***	***Injured/Month***
2019	5220	1567	6046	8335	870	261.2	1007.7
2018	4426	1560	5936	7037	738	260	989.3
2017	5836	1929	7217	9006	973	321.5	1202.8
2016	7101	2636	8865	9810	1183.5	439.3	1477.5
2015	6916	2808	8946	9041	1152.7	468	1491

Monitoring, analyzing, and improving driver’s behavior reduces traffic collisions and enhances road safety. Smartphone-based telematics systems are receiving more attention due to expanding sensing capabilities (accelerometers, magnetometers, GPS, etc.) and the popularity of mobile devices. Smartphone sensing capabilities have given a boost to the collection of high accuracy, real time, and high velocity data.

Several approaches have been investigated for the detection and identification of driver’s behavior. The conventional approach applies time series forecasting analysis methods (AR, MA, ARMA, ARIMA, and VARMA) for predicting the driver’s behavior assuming that data are stationary and ergodic,. Otherwise, data preprocessing is mandatory. Rule-based detection algorithms are extensively used to classify unseen data. However, they cannot distinguish the temporal relationships between timestamps, periodic, or seasonal irregularities. Deep learning approaches have succeeded to mine dynamical characteristics of time series. However, neural networks have some challenges that make them unsuitable for many classification tasks. The selection of efficient architectures, the optimization of hyper-parameters, the slow training, and the limited labeled data explains some of these challenges. Hybrid techniques can greatly increase prediction accuracy and get over single model limits for improved performance by integrating various models. Fusion and attention mechanisms were found to be more suitable for the analysis of time series sensor data. Transfer learning handles time-varying issues in time series data of various operational vehicle subsystems. Sharing the knowledge gained from simulating, the operational behaviors of various vehicle subsystems can also be helpful in modeling and predicting the behavior of other vehicles belonging to the same family [3].

Extracting vehicle’s maneuver represents a principal characteristic of driver’s behavior identification. A recent approach for extracting maneuvers from high-frequency telematics is through time series motifs detection algorithms [4]. The analysis of maneuvers using motif detection in telematics data is preferred over ML and DL and is considered a promising field of research that has not yet been fully investigated.

This work reviews the latest techniques for classifying driver behavior in time series data and summarizes the pros and cons of the different categories. The rest of this paper is organized as follows. Section II summarizes the definitions and approaches of time-series data analysis. Section III discusses the application of time series analysis techniques for driver's behavior classification. The application of motif finding algorithms for driver behavior detection, is discussed in section IV. The dataset section is cleared in section V. the discussion and recommendation are presented in Section VI. Finally, Conclusion in section VII

2. Time Series Data Analysis (TSDA)

2.1 Introduction [5 - 6]

In the field of the Internet of Things, where millions of linked devices have been deployed and where real-time data collection and processing are anticipated. It has become more and more important to be able to process time series data. Time-Series Data, also known as Time-Stamped Data, is an ordered set of measurements often obtained at regular intervals that may be divided into four categories: univariate, multivariate, tensor fields, and multifield. [7]

Time-series data processing and analysis are crucial steps for comprehending the data's features and collecting insightful statistics that support the investigation of the underlying processes, analysis, tracking, and display of this kind of data.

The two assumptions under which the theory of time-series data processing is built are Stationarity and Ergodicity [8]. However, time-series analysis can be applied where the series are seasonally stationary or non-stationary. Non-stationary data must be converted into stationary data using detrending and differencing in order to obtain repeatable and accurate results [8].

The analysis of time-series data [9] is strictly important for understanding the behavior of the data and obtaining meaningful information about the underlying processes. The cost needed to comprehend or spot abnormalities in time series data increases as their volume rises. An anomaly detection system can intelligently warn if something goes wrong. The classical analysis of time series data passes through several phases [10] as follows:

- Gathering and cleaning the data.
- Creating a visual representation that compares time to a critical element.
- Observes the series' stationary nature.
- Making charts to comprehend the nature of it.
- Building the models with AR, MA, ARMA, and ARIMA
- Drawing conclusions from a prediction

Linear autoregressive time series models are essential for modern stationary time series data analysis. They also provide useful foundation information for studying non-linear models [11].

The traditional time-series analysis pipeline has three factors: time-series data, similarity metrics & feature extraction, and time-series clustering [7]. Feature extraction is a type of dimension reduction that lowers the computational cost of working, with highly dimensional data, and increases clustering accuracy. The authors in [7] list a number of feature extraction techniques as a way to reduce time-series dimensionality.

Approaches for analyzing time series may be categorized into two categories: frequency-domain methods (such as spectral analysis and wavelet analysis) and time-domain methods [13] (Autocorrelation and cross-correlation analysis). However, there are two further categories of time series analysis techniques: parametric and non-parametric. Both linear and nonlinear time series models are possible. The goal of nonlinear time series analysis is to find structural breaks and model the behavior of the data before, after, and between breakdowns. The VARMA model, which is similar to the ARIMA model at univariate timescales, is the most widely used model for multivariable time series [13].

Time-series analysis applies Rule-Based Machine Learning (RBML) Models [14 - 16] to identify specific patterns. In unsupervised techniques, the raw data are first preprocessed, features are extracted and mined. Then clustering is applied for identifying specific patterns. In other words, time-series abstraction is carried out in three steps: preprocessing, discretization, and temporal aggregation [17] to provide a homogenous and condensed representation of the heterogeneous multivariate time-series.

Supervised learning techniques have found diverse applications for time series analysis. However, data labeling is very time consuming and costly. Additionally, time-series data has demonstrated a wide range of pertinent traits, attributes, complexity, and temporal scales. To get around these issues, a deep learning technique can be developed to separate the data manifolds, enabling a clustering technique to work with learnt features rather than raw data. Traditional clustering techniques have limited performance as dimensionality grows. Dealing with high-level representation has advantages that help clustering jobs be accomplished.

Deep clustering allows a deep neural network to locate hypothetical representative centers for scattered data by extracting common patterns in lower-dimensional space. Deep clustering approaches for picture datasets have been developed through efforts in the field of computer vision. Unsupervised models include deep auto-encoders (DAEs) and deep convolutional auto-encoders (DCAEs). Deep learning features were used to provide an abstracted latent representation for clustering analysis in these models. [12].

Available tools for time series data analysis, and few recent research applying these tools are summarized in Table 2.

2.2 Time Series Analysis Restrictions [9]

Time series data have some restrictions which must be considered during analysis, including:

- Like other models, TSA does not support the missing values.
- The relationships between the data points must be linear.
- Data transformations must be done, which makes them pricey.
- Most models operate on data that is only one variable.

Table 2. Time Series Data Analysis Models

1-Traditional Models [11]	2 – RBM Learning Models [18-20]	3-Deep Learning Models [21-33]	4 - Other Models & Methods [34-38]
AR, MA, ARMA ARIMA [11] VARMA [13]	Artificial neural networks	Convolutional Neural Networks [31, 32] Smoothed-CNN (S-CNN) [31] Multi-Channels Deep (MC-DCNN) [32]	Queuing theory analysis [34]
Fourier Transform DFT STFT	Support vector machines	Recurrent Neural Networks [24] • Including Elman [33] • Long-Short Term Memory • Gated Recurrent Units, and Bidirectional Networks	Control chart [35] • Shewhart individuals control chart • CUSUM chart • EWMA chart
Discrete, continuous or mixed spectra	Fuzzy logic	.	Nonlinear mixed-effects modeling
PCA Empirical orthogonal function analysis	Gaussian process	AutoEncoders [25]	Dynamic time warping [36]
Singular spectrum analysis	Genetic Programming	Deep belief network [28]	Dynamic Bayesian network [37]
Structural models	Gene expression programming	Sparse Functional Multilayer Perceptron (SFMLP) [26]	Time-frequency analysis techniques: • Fast Fourier transform • Continuous wavelet transform
General State Space Models	Hidden Markov model	Deep Graph-Evolution Learning [27]	Chaotic analysis [38] • Correlation dimension • Recurrence plots • Recurrence quantification analysis • Lyapunov exponents • Entropy encoding
Unobserved Components Models	Multi expression programming	N-BEATS (ElementAI) DeepAR (Amazon) Spacetimeformer [60] Temporal Fusion Transformer [52]	

3 . Time-Series Data Analysis for Driver Behavior Identification

3.1. Introduction

Detecting driver's behavior represents an interesting research area, as it can lead to decreasing the probability of traffic accidents caused by bad driving habits. The last few years have witnessed an excessive number of researches for detecting driver's behavior; starting by applying the classical time series prediction techniques, passing through rule-based and deep learning classification algorithms, and following by hybrid techniques.

The prediction techniques of time-series data common parametric approaches such as binary and ordinal logistic regression models and common non-parametric techniques such as Classification and Regression Trees and Multivariate daptive Regression Splines.The driver's behaviour has been extensively modelled and irregularities have been detected while taking into account temporal dependencies using the autoregressive integrated moving average (ARIMA) method [39].

Rule-based detection algorithms are extensively used to classify unseen data accurately, if sufficient amounts of labelled data are present. Even without labelled training data, new, unidentified sorts of anomalies may still be found .Additionally, since they are unable to understand the temporal connections between timestamps, it would be challenging to identify periodic or seasonal irregularities [40], [16].

Deep learning approaches have been applied to classify time series data. LSTMs and CNNs have succeeded to mine dynamical characteristics of time series. However, neural networks have some challenges that make them unsuitable for many classification tasks including the selection of efficient architectures, optimizing the hyper-parameters, slow training and limited labeled data [22].

Several Deep unsupervised learning approaches, including self-organizing map (SOM), deep autoencoders, and partitive clustering algorithms, were successfully applied for detecting drivers' behavior [41].

The peculiarities of different deep learning algorithms; supervised (Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) – Long Short-Term Memory (LSTM)), and unsupervised (Auto-Encoders (AE), Self-Organizing Map (SOM)), are summarized in [17].

In time series analysis, integrating hybrid linear/nonlinear techniques has received a lot of interest. [42] suggested a hybrid ARIMA-WANN strategy that combines the Wavelet Autoencoder Neural Network and the ARIMA model to effectively identify abnormalities in huge time series data sets of vehicle operating data.

For better performance, it is logical to combine different models through hybrid approaches that can considerably improve the prediction accuracy and overcome the limitations of single models [42]. Fusing CNN and RNN with an attention mechanism, was found to be more suitable for time series CAN-BUS sensor data [43 - 45].

One-Class Support Vector Machine, Autoencoder scheme, and Long Short Term Memory network are all components of a multivariate time series model that has achieved significant success in predicting vehicle behavior and detecting unsafe conditions. [46].

In addition to the above research studies, several data analyzers and simulators were implanted for studying the drivers behavior from datasets gathered for this purpose. A Real-time Analyzer for driver’s behavior datasets was implemented by [47]. A Driving Data analysis tool was created by INVERS [48] to detect unwanted driving behavior. [157] implemented a drive pattern analyzer using smartphones. In addition, Simulators were implemented to create data suitable for driver behavior analysis [49,50].

3.2 Traditional and ML Time-Series Data Analysis Techniques for Driver Behavior Identification

3.2.1 Traditional (Rule-based) Machine Learning Models

Security data set [14], HciLab [112], and UAH-DriveSet [98] datasets were used by the authors of [14] to put up an evaluation system for comparing the effectiveness of current driver behaviour identification systems.They looked for methods to increase their efficiency and attain high recognition accuracy while maximising the number of features (training dataset size, and identification time).The machine learning techniques taken into account for the classification problem include Decision Tree, Random Forest, Extra Trees, KNN, SVM, Gradient Boosting, AdaBoost based on Decision Tree, and multi-layer perception (MLP).Their method involves first segmenting the datasets into small chunks, then applying classification algorithms to an increasing number of those chunks until the identification score meets a predetermined level of satisfaction (preferably 100%).Every minute, the driver identification process was carried out, resulting in at least 60 records per characteristic in the identification task.To compare the various classification algorithms, they used a cross-validation of 10 folds. The Random Forest and Extra Trees algorithms were found to perform better than any other method examined in this article. To properly comprehend this issue, more investigation is necessary since the case studies in this work only include a small number of drivers.

The researchers in [51] presented a rule-based machine learning technique utilizing a sequential covering algorithm To categorize driving maneuvers from time-series data obtained from the sensor fusion time series dataset. The influence of each rule is evaluated using the metrics of coverage and accuracy in the sequential covering method. These metrics represent the quantity of covered and correctly identified examples in a maneuver class, respectively. Only the important constraints are used to create each maneuver class's final rule set. In this method, the rules are learned without supervision, and only those that perform the best are added to the ruleset. When compared to conventional machine learning and deep learning methodologies, which frequently call for a larger dataset, more sophisticated computations, and longer processing times, the proposed system's use is advantageous.

Time series data can be classified using deep learning (DL) models; however, it is unclear how these algorithms handle the data in between these classification steps. The Rule-Based Machine Learning (RBML) approach, on the other hand, learns rules using domain experts' understanding of a problem area and expresses rules logically. Therefore, it's crucial to comprehend a problem domain's fundamental problems when creating an RBML model. However, it can be quite difficult to categorize time series data using rules [52].

[53] Utilized four methods RFT, Gaussian Naive Bayes, SVM, and Logistic Regression, and determined the following performance metrics: In every confusion matrix, accuracy, precision, recall, mistake rates, and specificity are all considered.

An important metric for making a choice on the optimal model is the F1 score, which is the harmonic mean of accuracy and recall. according to SVM, the F1 score of Random Forest was the greatest for Gaussian Naïve Bayes. All of the algorithms' specificities were discovered to be between 82% and 90% or higher. Additionally, this is a crucial element. Each method used in this investigation produced incredibly precise results.

3.2.2 Deep Learning Models:

Time series data can be classified using deep learning (DL) models; however, it is unclear how these algorithms handle the data in between these classification steps. In the “Stacked-LSTM” architecture used by [60], two LSTM memory cell layers are used, each with 100 hidden neurons. The first LSTM layer is fed by a time-series window S with size 64 feature vectors of 9 fused sensor data, recorded using a smartphone's internal sensors, at each time step. The proposed model was evaluated on UAH-DriveSet [98], obtaining cutting-edge outcomes with greater true positive rates and lower false positive rates compared to the baseline technique. evaluating how the suggested Stacked-LSTM model performs in comparison to a variety of widely-used classification algorithms in studies that classify and analyse driving behaviour ,they improved by more than 10% over the closest comparative techniques to get an F1-measure score of 91%.

Deep learning (DL) models can be used to categorize time series data, but their early processing stages remain unknown. In the “Stacked-LSTM” architecture used by [60], two LSTM memory cell layers are used, each with 100 hidden neurons. The first LSTM layer is fed by a time-series window S with size 64 feature vectors of 9 fused sensor data, recorded using a smartphone's internal sensors, at each time step. The proposed model was evaluated on UAH-DriveSet [98], achieving state-of-the-art results with higher true positive rate as well as lower false positive rate in comparison to the baseline approach. When the suggested Stacked-LSTM model was compared to a variety of standard classification algorithms used in driving behavior classification and analysis studies, it earned an F1-measure score of 91%, an increase of more than 10% over the closest comparative methods.

Authors in [54] presented an approach for classifying driving behavior, based on stacked LSTM Recurrent Neural Networks. Using a smartphone's inbuilt sensors to record data from nine distinct sensors. Capturing nine different sensor data using the internal sensors of a smart phone, they formulated the driving behavior problem as time-series classification problem (normal, aggressive or drowsy). They tested their suggested Stacked-LSTM model using UAH-DriveSet [98] and obtained cutting-edge results. They compared the performance of their proposal against a number of common driving behavior classification and analysis algorithms and found that their approach achieved F1-measure score improvement of more than 10% over the closest compared approaches.

On the other hand, [55] used their Stacked-LSTM, a customized form of LSTM architecture made up of two LSTM cell layers and arranged as a many-to-one structure. With 120 hidden neurons and the ReLU activation function, both LSTM layers are set up. Preventing overfitting by employing the L2 regularization approach in both layers. They determined that a window size of 16 time-steps is the best since it produced the maximum accuracy after conducting several trials with various window sizes.

The authors in [56] used the three layers of the "standard LSTM neural network": an input layer, a recurrent hidden layer, and an output layer. The suggested LSTM network structure is tested and trained using drivers' driving signals over a short period of time. The size of the window (T), according to experiments, has a significant impact on the categorization outcomes. Sample data are insufficient to differentiate between drivers when the T is too tiny. On the other hand, a big T results in unrealistic signal shapes because of the training road's repeating character. Therefore, a 30-second window is chosen that gives each sample a respectable quantity of information concerning driving behaviors. Additionally, adjusting the window for a 15-second interval has improved training results. The suggested method is capable of dynamically identifying the driving behaviors of real drivers in discrete time.

The researchers in [95], in their trials to detect drivers’ behavior time series patterns using their own dataset, during the driving lesson, RNN (GRU, LSTM) algorithms were used to detect patterns. The algorithm is unaffected by road factors, vehicle characteristics, location, or driver conduct. They found that (120 seconds of GPS data) is adequate to determine a driver's behavior. According to their own dataset, their results have high accuracy. However, this proposal should be validated using other datasets to ensure its real-world applicability.

The authors in [57] proposed an “FCN-LSTM” architecture, The FCN network is made up of numerous temporal convolutional layers and serves as the model's feature extraction module.Using a sliding window with a defined length, the data were split into overlapping parts before being injected into the algorithm. The proposed model was evaluated on the public UAH-DriveSet dataset [98], utilizing the right amount and variety of features, the best up-sampling ratio for the signals, and the optimal window size.When performed with the above dataset, the suggested model improved the maximum known F1-score of 91% by 8.49%, according to comparisons with cutting-edge models.

The authods in [58], applied a stacked LSTM network with attention to identify risk of road accidents and compared it to both stacked LSTM and MLP models to illustrate the significantly positive effect of applying attention on simulation resultsThey collected a large dataset of driving data, and conducted an experiment with eight driving scenarios. Results show that adding attention to the stacked LSTM model decreased the training and testing errors, diminished overfitting, and reduction in computational expenses.

The authors in [59] proposed a driving behavior detection method based on Soft Thresholding and Temporal Convolutional Network (S-TCN). To increase the model's stability and accuracy, they combined the TCN with soft thresholding. The suggested approach is thoroughly tested on four real-world public data sets. The suggested model outperforms the best state-of-the-art baselines by 2.24%, according to experimental data.

Table 3 presents examples from the recent researches published in the different categories of vehicle’driver behavior.

3. MOTIF DISCOVERY FOR DRIVER BEHAVIOR DETECTION

4.1. Time Series Data Analysis Techniques Using Motifs

Time series data can be found everywhere measuring any phenomena over time generates time series data that may be collected using appropriate sensors. Time series data representations (TSDR) are intended to: reduce the dimensionality of the raw time series for reducing the computational burden, If the aim is to do classification, clustering, or anomaly detection, reshape the time series to obtain it in the required format, summarize the information contained in the time series, remove noise, and manage distortions, or highlight discriminative or distinctive information [87].

TSDRs change depending on the type of data summarized and the structure of the representation. TSDR can be Time-based, Feature-based, or Motif-based. Time-based visualizations create a time series from the complete raw of time series. The raw time series is converted into a scalar or a set of scalars through feature-based representations. Motif-based representations generate a subsequence based on desired attributes retrieved from the original time series. Motif-based representations are often dataset-focused; motifs are found at the dataset size. Motif-based representations are more often dataset focused; motifs are discovered at dataset scale. Motifs are subsequences of the raw data, previously unknown; they could be recurrent, infrequent/surprising, or discriminant. The frequency of occurrence of surprising motifs significantly differs from that expected by chance [87].

Table 3 Algorithms for Detection of Drivers' behavior: Examples of Related Research.

Rule-Based ML Mode [14 - 16]	Supervised Models [61], [42]	Auto-Regressive, Moving Average (ARIMA) Models	[42]
		K-Nearest Neighbor (KNN)	[14]
		Logistic Regression (LR)	[14], [62]
		Decision Trees (DT)	[14]
		Regression Trees (RT)	[63, 64]
		Random Forests (RF)	[65, 66], [16]
		Hidden Markov Models (HMM)	[15], [62], [68,69]
		Neural Networks (NN)	[46], [15], [66, 67], [6
		Multilayer Perceptron (MLP)	[14], [52], [70]
		Support Vector Machines (SVM)	[62], [14]
		Gradient Boosting	[14]
		Gaussian Mixture Model (GMM),	[15], [16], [72]
		fuzzy control theory,	[15]
		Other Methods (Entropy,	[71]
		Hybrid Models (Hybrid ARIMA-WNN)	[42]
	Unsupervised Models [40]	Clustering	[73 - 76]
		Principal Components Analysis (PCA)	[39]
		Structural Topic Modeling (STM)	[39]
		Unsupervised Sequential Covering	[40]
		Representation Learning	[77]
		Unsupervised sequential covering approach	[40]
Deep Learning Models [17],[61], [42], [72]	Supervised Models [42]	Deep ARIMA	[42]
		Convolutional Neural Networks (CNN)	[70 - 80]
		Recurrent Neural Networks (RNN)	[17],[80-82]
		Long Short Term Memory (LSTM)	[17], [80]
		Hybrid Models (Hybrid ARIMA-MLPNN and ARIMA-WNN	[42]
		Temporal Convolutional Network And Soft Thresholding	[65]
	Unsupervised Models [41]	Auto-Encoders (AE)	[17], [83], [41]
		Stacked Autoencoder (SDAE)	[84]
		Representation Learning	[85], [77]
		Self-Organizing Map (SOM)	[41]
Fuzed & Hybrid Mode		Multi-Source Data Fusion	[43]
		Fusing CNN and RNN with an attention mechanism	[44]
		(CNN-LSTM)	[45]
Transfer Learning		Transfer Learning methods for temporal data	[86]
Offline Methods	Analyzers	Real time Analyzer, INVERS Ansayzer	[47 - 48]
	Simulators	Driver Behavior Simulator, Simulator Driving Tool	[49 - 50]

The process of obtaining a motif or a group of motifs from time series data is known as motif discovery. There are four categories of motif discovery algorithms: enumerative, probabilistic, combinatorial, and nature inspired. Each category includes several subcategories [88].

Enumerative techniques are exhaustive search techniques; they are the only techniques that ensure finding all motifs. However, they are slow, and require a lot of user defined parameters. Probability-based approaches overcome drawbacks of enumerative approaches, but they are complex and can't find all motifs. Natured inspired approaches combine the main features of the enumerative and probability-based approaches. They can deal with the big data and long motifs. Combinatorial approaches are hybrid algorithms that mix multiple algorithms for better accuracy; their common features are the flexibility of objective function [88].

Despite the reality that motifs are associated to time-series segments, the exact meaning of the most meaningful motifs different between researchers and areas of usage. According to [89], there are two basic definitions: similarity-based and support-based. The motif in the similarity-based can be arranged according to the similarity of the parts that correspond to the motif (highly similar motifs). Support-based ordering sorts of motifs according to the number of repetitions the motif appears in the time series (highly frequent motifs).

Since the formalization of temporal motifs in 2002, various researchers have employed them in a wide range of applications. Because of the difficulty of motif recognition algorithms, many of approximation techniques for detecting motifs have been proposed.

The authors in [90] suggested a workable efficient approach for finding time series motifs in massive datasets that is up to a factor of three quicker than brute-force search. They also demonstrated that the suggested technique is sufficiently quick to be employed as a subroutine in higher-level data mining algorithms for classification, detection, and summarization. The defining of motif length, managing data streaming, motifs with varying lengths, computation time and discovery of unknown motifs are all common difficulties in motif discovery algorithms. Based on the application, motif discovery algorithms vary in how they solve challenges. These algorithms may be adjusted to discover accurate or approximate patterns, as well as motifs of variable or constant length. They can handle multivariable or simple time series and perform in either on-line or off-line [91].

In recent years, there has been a lot of interest in finding variable-length motifs. To identify variable-length motifs, the existing situation algorithm employs a fixed-length motif discovery algorithm as a subroutine to identify it.

[92] presented a Hierarchical-based Motif Enumeration (HIME) algorithm as an estimated technique for detecting variable-length motifs with a wide sampling spectrum in million-scale time series. The authors found, through experimentation, that the scalability of the proposed algorithm is significantly better than that of the state-of-the art algorithm. They demonstrated that HIME can efficiently detect meaningful variable-length motifs in long, real world time series.

The authors of [93] developed a Motif-Aware State Assignment method (MASA) to discover common motifs in noisy time series data, formulated as a large optimization problem, and solved using an expectation-maximization approach. The suggested approach works well even when there is noise in the input data and can handle very big datasets. Studies using fictitious data revealed that MASA performs up to 38.2% better than cutting-edge baselines.

The distance utilized to compare time-series segments is another critical aspect of motif finding. The Euclidean distance or the Dynamic Time Warping distance is used in most articles. While the first method is highly efficient and allows for quick segment comparisons, the second allows for segment comparisons with varying lengths and time distortions by determining the ideal time alignment of the two time-series being compared.

[94] Proposed an approach that enables the flexible comparison of motifs of varying durations. They suggested, in particular, using the Symbolic Aggregate Approximation (SAX) representation to generate a discrete list of words and aggregate recurring terms inside the same comment.

Recently, the area of time-series motif discovery has received a lot of attention from the data mining community. In spite that myriad researchers have investigated time series motif detection in many applications including computational biology, genetics, medicine, entomology, weather prediction, seismology, entertainment, etc., few trials have been carried out to investigate the use of time series motifs for studying vehicles driver's behavior; the key in many real-world situations that can help reduce traffic accidents all over the world.

4.2. Motifs for Driver Behavior Detection

Recently, the area of time-series motif discovery has received a lot of attention from the data mining community. In spite that myriad researchers have investigated the time series motif detection algorithms in many applications including computational biology, genetics, medicine, entomology, weather prediction, seismology, entertainment, etc., few trials have been carried out to investigate the use of time series motifs for studying vehicles driver's behavior.

The Extended Motif Discovery (EMD) algorithm is a time-series motif identification method that can identify subsequences of varying lengths inside the same motif. The EMD method, introduced in [94], is a popular time - series data motif finding approach based on the concept of minimum description length. One distinguishing aspect of the EMD method is that it can identify the appropriate length of the motif and therefore does not need the length of the motif as a user-supplied input. Another intriguing aspect of the EMD is that the lengths of each iteration of a motif can change slightly from one another. As a result, they calculated the distance between the motif repetitions using dynamic time warping (DTW) distance, resulting in a high computational complexity and complexity of practical implementation.

The authors of [96] presented an effective EMD algorithm solution in that they utilized homothetic conversion to turn all motif occurrences of varying lengths to the same length. to make calculating Euclidean distances between them easier. This change significantly speeds up the execution of the EMD and makes it easy to apply. The updated EMD implementation approach for time series motif finding is illustrated by experiments on 7 real-world time series datasets.

The authors in [97] presented an approach for identifying maneuvers from vehicle telematics data, through motif detection in time-series. They used a modified version of the Extended Motif Discovery (EMD) method [95] that was applied to the UAH-DriveSet [98]. The first experiment attempted to detect acceleration and brakes from longitudinal acceleration time series, and the second attempted to recognize turns from lateral acceleration time series. They noticed that the updated EMD algorithm successfully extract complicated maneuvers including lane changes and overtaking movements.

The authors in [4] proposed a system (TripMD) that extracts relevant driving patterns from a set of trips. They used Extended Motif Discovery (EMD) algorithm to find their motifs, they also employed the DBSCAN algorithm. To evaluate the applicability of TripMD to real tasks, they used the UAH-DriveSet [98]. The three behaviors identified in the dataset (normal, aggressive, and drowsy) were identified.

4.3. Evaluation Metrics

According to [113], the performance evaluation of the time series analysis methods is examined from two viewpoints: simplicity of the model, and model accuracy [114].

Time series metrics or features that can be used for time series classification can be categorized as linear or nonlinear. In each case, these metrics depends on the type of data; whether Univariate, Bivariate or Multivariate [115]. However, Similarity measures, including Cross-correlation, Dynamic Time Warping, Hidden Markov Models, Distance Measures, Total correlation, are mostly used.

First-order temporal correlation coefficient and Euclidean distance for standard values are two often used similarity metrics in time series data analysis. The Euclidean distance metric is extremely susceptible to noise and distortion and is unable to deal with one of the components being compressed or extended. As a result, this method is unreliable, especially when determining how similar two time-series with differing slightly different are [116-118] published a detailed study of time-series measurements, dividing them into four groups:

- Lockstep actions such as Euclidean distance and Manhattan distance.
- Elastic measures, such as dynamic temporal warping (DTW) and longest common subsequence (LCS).
- Measurements based on patterns, such as the spatial assembling distance (SpADe).
- Measurements with thresholds, such as threshold query-based similarity search (TQuEST).

To get around some Euclidean distance restrictions, such as non-linear distortions, Dynamic Time Warping (DTW) is proposed. The principle behind DTW is to align (warp) the time-series before calculating the distance, thus the two time-series don't have to be the same length. But DTW could match two temporal points that have very distinct local structures. The alignment technique may be made better, for example, by using shape dynamic temporal warping, which takes point-wise local structural information into account [119]. When working with huge datasets, DTW does not scale well due to its quadratic time complexity. Boundary, monotonicity, and continuity constraints are only a few of the local restrictions that apply to DTW. Common misconceptions regarding DTW include the ideas that it is too slow to be effective and that the size of the warping window is not very important [120].

5. Datasets

Bench-mark datasets are essential for evaluating ML & DL algorithms. For driver behavior detection, several datasets are available, including:

- **UAH-DriveSet [**98]

The UAH-DriveSet provides more than 500 minutes of naturalistic driving, simulating different driver behaviors (aggressive, drowsy or normal) on two different types of roads (highway and secondary). A variety of drivers in various situations and actions captured UAH-DriveSet using the mobile application DriveSafe. *UAH-DriveSet* includes both raw data from the GPS and accelerometer, and processed data from videos representing maneuver recognition and driving style estimation.

- **Kaggle Driving Behavior Dataset [99]**

On the basis of accelerometer and gyroscope data, the Kaggle Driving Behavior Dataset is intended for modeling unsafe driving behaviors. Three distinct drivers who were 27, 28, and 37 years old collected the data on Hyundai i20, Ford Fiesta 1.4, and Ford Fiesta 1.25 vehicles. The actions were quick acceleration, sharp breaking, sharp turns to the right and left.

- **HDD (Honda Research Institute Driving Dataset) [100]**

The Honda Research Institute Driving Dataset (HDD) consists of 104 hours of actual human driving in the San Francisco Bay Area that were recorded using an instrumented car with various sensors. The collection includes 104 hours of driving movies shot by real people that were timed to CAN bus signals, GPS, and IMU data. There are 137 driving sessions total.

- **Mendeley Data: Smartphone Sensor Dataset [101]**

More than 15,000 sensor measurements for a one-way trip of 5–17 kilometres are included in the data collection. Simulated real-world driving scenarios were used to extract data from smartphone sensors. Driving activities were conducted using a range of cars and in a variety of road and traffic conditions in order to carry out the events and gather data.

- **Driving Behavior Dataset Version 2 [102]**

This Dataset is based on accelerometer and gyroscope data, built in Ford Figo 1.2, Maruti Suzuki Swift VXI, Tata Nexon XMS. 3 different drivers with the ages between 35-40 yrs, engaged in driving activities of Normal, Aggressive, and Risky Driver Behaviors. Smartphone Redmi 4, MI with: Accelerometer, Gyroscope Sensors was used.

- **Drive&Act [103]**

Drive&Act is a multi-modal dataset that self-driving cars utilize to detect precise driver behavior. The Drive&Act benchmark dataset contains 12 hours and more than 9.6 million frames of people engaged in distracting activities using a standardized multiview camera system with five perspectives, including color, infrared, depth, and 3D body position information.

- **IEEE Vehicle Driving Behavior Dataset [104]**

Two text files, titled "RAW DataSet.txt" and "Augmented dataset.txt," make up this dataset. With the use of an accelerometer and gyroscope, 1,032 events—including acceleration, collisions, left and right turns, and routine driving—were recorded. Each event's 300 samples were then enlarged. In order to create a balanced dataset of 10,000 samples from the initial 1032 imbalanced data, where the quantity of each event sample was equal.

- **Driver Behavior Dataset [105]**

This dataset includes a collection of smartphone sensor readings that simulate common driving actions including braking, accelerating, turning, and changing lanes. While a driver performed specific driving actions, sensor data from a smartphone was recorded using an Android application. Four 11-minute-long automobile excursions (2011 Honda Civic) were used to obtain the data. The Motorola XT1058 smartphone, running Android 5.1, was mounted to the windscreen of the vehicle. The driving events were conducted by two drivers with a combined total of more than 15 years on the road.

- **DBNet [106]**

A big dataset for studying driving behavior is called DBNet. It gathers data on 1000 KM of actual driving using aligned video, point clouds, GPS, and driver behavior (velocity and wheel).

- **DriverMVT [107]**

While individuals were traveling to or from work, DriverMVT (In-Cabin Dataset for Driver Monitoring) was recorded in a real-world setting. 1506 footage of drivers inside of cars make up the dataset. The Samsung Galaxy S10 and S20 cameras, as well as a USB camera made by ELP using an OV7725 sensor, were used to acquire the data. Videos were captured for the smartphones at a resolution of 1080 x 1920 and a frame rate of 60 fps. The dataset was collected from 9 drivers of various ages and genders, with a total of 5119k frames and over 36 hours of time spent driving under various light and speed circumstances.

- **The "KITTI" dataset [108]**

The "T-Drive trajectory dataset" from the KITTI dataset shows a path of 10.357 taxis over the period among one week. The total amount of points in this dataset is around 15 million, and the total distance is nine million kilometers.

- **The HCRL dataset [109, 110]**

The HCRL dataset comprises a 46-kilometer trip with 10 drivers across Seoul and the surrounding areas, taking around 23 hours to complete from Korea University to the SANGAM World Cup Stadium. Fabio Martinelli et al. provide an in-depth summary of the dataset in [D13].

- **The "Warrigal" dataset [111]**

The "Warrigal" dataset is a big, comprehensive dataset that was gathered from the experiences of big trucks and little 4WD cars in an industrial setting. The data was gathered over the course of three years by a fleet of 13 vehicles operating in a surface mine. The information includes details about the vehicles' state (such as their position, speed, and heading), as well as summaries of their radio contacts with one another. The data spans three years at a resolution of 1 Hertz. This dataset has already been used to examine wireless network antenna failure, driver purpose inference, protection analysis, and map construction.

- **Security dataset [14]**

KIA Motors Corporation's most recent model was used in the data collecting in South Korea. Ten drivers took part in the experiment, which was conducted on a 23 km total of four tracks of three different types—city way, highway, and parking area. 94,401 records were recorded on average per second, resulting in a 16.7 MB dataset. Data were gathered from the CAN bus of the car using the On Board Diagnostics 2 (OBD-II) and Carbigs systems (OBD-II scanner). The Electronic Control Unit (ECU) of the used car controls several measurement and control sensors. The OBD-II system was used to measure a total of 51 characteristics.

- **HciLab dataset [112]**

The HciLab Driving dataset, which is available to the general public, is 450 MB in size and contains 2.5 million samples. Information concerning GPS, brightness, acceleration, physiological data, and video rating data are all included in the anonymized file. The driving session produced three distinct data sets. Initially, three sensors linked to the participant were used to record the physiological status of the driver. Second, an Android smartphone was used to gather context data (Google Nexus S). The driving situation of the road and the driver's perspective were captured using cameras from Logitech QuickCam Pro 9000 and Creative VF0610 Live.

The HciLab Driving dataset, publicly available, Apart from these datasets, some researchers have created and used their own collected data for detecting the driver's behavior. The authors in [56] used IPG's TruckMaker vehicle dynamics simulation software to develop an exact physical model of a commercial Ford Truck using the vehicle's true parameters and qualities. Each driver is simulated on both training and test roads using the vehicle model and five various trailer loads. Similarly, the authors in [43] used the data collected by themselves (3,120,578 pieces of vehicle kinematic data and 905.5 GB of video clips). The vehicle kinematic data contained 12 different features about the vehicle position, speed, acceleration, and attitude The video clip data include the information from outside ambient videos and driver videos. In [17], Data is retrieved through a mobile app that is already installed on both iPhone and Android cellphones. While commuting, the software is still functioning in the mobile operating system without any user involvement. The application uses GPS cameras, accelerometers, and gyroscopes to collect raw data from cellphones. Data is communicated to the server via cloud-based services, where it is anonymously stored for further study following automatic identification at the end of the journey

6. Discussion and Recommendations

Recently, there has been a lot of interest to study effectively extracting newly undiscovered, recurrent patterns from time-series data. For many time-series data mining jobs, these patterns, or "motifs," are particularly helpful. The literature is full of research investigating motif discovery approaches applied to various areas as telecommunication, medicine, web, motion-capture, and sensor networks.

In order to address peculiarities in human behavior and the ensuing safety concerns, driver behavior is presently the subject of substantial research.

Driver behavior is now the topic of substantial study, which is being performed to overcome human behavior idiosyncrasies and the related safety difficulties. Monitoring, analyzing, and improving driver’s behavior reduces traffic collisions and enhances road safety.

Smartphone-based telematics systems are receiving more attention due to expanding sensing capabilities and popularity of mobile devices, giving a boost to the collection of real time high velocity data.

Several strategies have been employed to detect and recognize the target driver’s behavior. However, handling of the smartphone's loudness, inexpensive inertial sensors built into smartphones, the vibrations caused by vehicle engine, in addition to the road conditions and other issues make it difficult to accurately detect anomalous driving when using a smartphone. As a result, cleaning out the numerous sounds in raw data is required to increase the accuracy and stability of driver behavior detection. The model used presents a second obstacle. To avoid overfitting, the sophisticated deep learning models need a large amount of training data. Real-time prediction on mobile phones requires a lot of time and storage, which the training and prediction process cannot provide [75].

The temporal feature of the time-series dataset implies the data processing in a well-defined time period, called “window”, (a sliding window). A drawback of these solutions is the rigidity of the window size. The way a trip is divided, if overlap is not used, may not be ideal because a window can potentially divide a single action and information may be lost. However, employing the maximum overlap (i.e., moving the window by one time step) has drawbacks of its own.

Motif-based clustering plays an important role to overcome the problem of window size. Motifs are subsequences of the raw data, previously unknown; they could be recurrent, infrequent/surprising, or discriminant. The frequency of occurrence of surprising motifs significantly differs from that expected by chance [113]. Motif-based representations produce a subsequence extracted from the raw time series based on desired properties. The majority of motif-based representations are dataset-focused, and motifs are found at dataset size. The process of extracting a theme or group of motifs from a time series of data is known as motif discovery. There are four categories of motif discovery algorithms: enumerative, probabilistic, combinatorial, and nature inspired. Each category includes several subcategories. [114].

Since the formalism of time series motifs in 2002, myriad of researchers has used them for diverse applications in different domains. The complexity of motif discovering algorithms has led to the proposal of dozen approximate algorithms to discover motifs.

Defining motif length, handling data streaming, handling motifs with varying lengths, temporal complexity, and finding obscure motifs are common issues with motif discovery algorithms. The approach taken by motif discovery algorithms varies depending on the application. These techniques can be modified to locate precise or approximative patterns or to recognize motifs with variable or constant lengths. They may run in an off-line or online mode and handle multivariate or univariate time series [117].

The distance applied while comparing time-series segments is a significant factor in motif finding. The Euclidean distance or the Dynamic Time Warping distance are typically used in academic articles. The second has the benefit of enabling the comparison of segments with varying lengths and temporal distortions by computing the ideal alignment in time of the two time-series that are being compared, but the first is extremely efficient and enables a quick comparison of segments.

Recently, the area of time-series motif discovery has received a lot of attention from the data mining community. In spite that myriad researchers have investigated time series motif detection in many applications including computational biology, genetics, medicine, entomology, weather prediction, seismology, entertainment, etc., few trials have been carried out to investigate the use of time series motifs for studying vehicles driver’s behavior.

Recommendations

- Gathering vast amounts of driving information from different vehicles and drivers will be of utmost relevance in the future.
- Typically, real-world data is incomplete, inconsistent, inaccurate (contains errors or outliers). Data preprocessing is necessary for cleaning, formatting, and organizing the raw data, making it suitable for subsequent processing.
- The biggest finding from the forecasting M-Competitions is that combinations of various methods have consistently outperformed any particular method. For better performance, combining different models through hybrid approaches can considerably improve the prediction accuracy and overcome the limitations of single models. Fusion and attention mechanisms were found to be more suitable for the analysis of time series sensor data.
- Myriad researchers have investigated time series motif detection in many applications including computational biology, genetics, entomology, weather prediction, seismology, entertainment, etc. However, few trials have been carried out to investigate the use of time series motifs for studying vehicles driver’s behavior.
- In the field of driving behavior analysis, motif identification appears to be a viable study area.
- Vehicle’s maneuver represents a principal characteristic of driver’s behavior identification. A fascinating area of study that has not yet been completely investigated is the analysis of maneuvers using motif identification in telematics data.
- Extracting maneuvers from high-frequency telematics through time series motifs detection algorithms is preferred over ML and DL approaches as it had the advantage of not requiring labels, which is extremely time-consuming to collect.

7. Conclusion

The analysis of time-series data is necessary for understanding the behavior of the data and discovering meaningful information about the underlying processes. As the volume of time series grows, the effort required to understand or detect anomalies in it becomes very costly.

Autoregressive time series models are central to time series data analysis. For short-run high frequency data, their results may be hard to beat. However, they are unsuitable for multivariate datasets, computationally expensive, and have poor performance for long term analysis. The most common traditional time-series analysis autoregressive models are AR, MA, ARMA, AARIMA, and SARIMA. In addition to these classical models, there are many complex models that can treat the shortages of the classical algorithms, including GARCH, Bayesian-based models, and VARMA.

Neural Networks Autoregression (NNAR) models are examples of neural network models that may be used with time series that employ lagged predictors and can handle features Rule-Based Machine Learning (RBML) models that learn rules from domain expert knowledge. When creating these models, it's crucial to comprehend the underlying problems in a certain problem area. However, it can be quite difficult to categorize time series data using rules.

The use of Deep Learning approaches for time-series data analysis meets plenty of challenges, including lack of training data, poor quality of data, data overfitting, or underfitting. Critical data challenges include Quality, Sparsity, and Integrity. In addition, the intermediate processing steps of Deep Learning models are unclear.

Training a machine learning model might be difficult when there is a lack of labeled data. You may overcome this obstacle in a number of ways, though. Transfer learning, active learning, self-supervised learning, semi-supervised learning, and weakly supervised learning are a few of the available methods. Machine learning models can generally be restricted by their accuracy, the kinds of issues they can address, and the caliber of the training data.

In comparison to single models (such ARIMA or LSTM), some authors have shown that the hybrid implementation is more effective at capturing both the linear and non-linear features of the datasets. According to their findings, the hybrid of ARIMA and NARNN (Nonlinear Auto-Regression Neural Network) outperformed one of the most popular time-series prediction techniques by about 35.3%. (ARIMA). These results underline that to show that the complexity of machine learning and deep learning methods for time series forecasting is adding skill to the prediction, it is necessary to analyze both classical approaches and their outcomes as a baseline.

References

1. World Health Organization. Global status report on road safety, 2022.
2. Central Agency for public mobilization and statistics, Egypt, https://capmas.gov.eg/HomePage.aspx
3. Sablons de Gélis, “Transfer learning techniques in time series analysis”, Master Thesis Report, KTH Royal Institute of Technology, Electrical Engineering and Computer Science, Stockholm, Sweden 2022.
4. Maria Ines Silva, Roberto Henriques, “TripMD: Driving patterns investigation via Motif Analysis”, Expert Systems with Applications, Volume 184, 2021. DOI:10.1016/J.ESWA.2021.115527
5. Heckert, N., Filliben, J , Croarkin, C., Hembree, B., Guthrie, W., Tobias, P. and Prinz, J., Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, MD (Accessed October 23, 2022).https://doi.org/10.18434/M32189
6. What Is Time Series Data? | Definition, Examples, Types and Uses. InfluxData, www.influxdata.com/what-is-time-series-data. Accessed 20 Nov. 2022. https://www.influxdata.com/what-is-time-series-data/
7. Alqahtani, A.; Ali, M.; Xie, X.; Jones, M.W, “Deep Time-Series Clustering: A Review", Electronics, 10(23):3001, 2021. https://doi.org/ 10.3390/electronics10233001
8. Rob J Hyndman and George Athanasopoulos, "Forecasting: Principles and Practice", oTexts, 3rd edition, May 2021].

9. Pandian, Shanthababu. “A Comprehensive Guide to Time Series Analysis - Analytics Vidhya.” Analytics Vidhya, 23 Oct. 2021, www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-to-time-series-analysis.
10. James Andrew Godwin, "Time Series Analysis", Towards Data Science, 2021. https://towardsdatascience.com/time-series-analysis-7138ec68754a
11. Raquel Prado and Mike West, “Time Series Modelling, Inference And Forecasting”, AMS, University of California, Santa Cruz ISDS, Duke Universit tsbook.dvi (hughchristensen.com)
12. Al-Qutt, M. M., Heba Khaled, Rania Elgohary, Hossam M. Faheem, Iyad A. Katib and Nayif Al-Johani. “Accelerating Motif Finding Problem Using Skip Brute-Force on CPUs and GPU’s Architectures” (2017).
13. Tamás Ferenti, "Biomedical applications of time series analysis", 2017 IEEE 30th Neumann Colloquium (NC), IEEE Xplore, Budapest, Hungary, 2017 DOI:
14. Saad Ezzini, Ismail Berrada and Mounir Ghogho,“Who is behind the wheel? Driver identification and fingerprinting”, J Big Data,pp 1-15, 2018. https://doi.org/10.1186/s40537-018-0118-7
15. Na Lin, Changfu Zong, Masayoshi Tomizuka, Pan Song, Zexing Zhang, Gang Li, "An Overview on Study of Identification of Driver Behavior Characteristics for Automotive Control", Mathematical Problems in Engineering, vol. 2014, Article ID 569109, 15 pages, 2014. https://doi.org/10.1155/2014/569109
16. Ferreira J Júnior, Carvalho E, Ferreira BV, de Souza C, Suhara Y, Pentland A, et al. (2017) Driver behavior profiling: An investigation with different smartphone sensors and machine learning. PLoS ONE 12(4): e0174959. https://doi.org/10.1371/journal.pone.0174959
17. Farid Talebloo, Emad A. Mohammed, Behrouz H. Far, “Dynamic and Systematic Survey of Deep Learning Approaches for Driving Behavior Analysis”, Computer Science, Machine Learning, Cornel University, arXiv:2109.08996, Top of Form 2021. https://doi.org/10.48550/arXiv.2109.08996
18. Pablo Ruiz,”Machine Learning Approaches for Time Series Data, Modeling time series with non conventional models”, 2019.Machine Learning Approaches For Time Series Data
19. Antonio Rafael, SabinoParmezanVinicius, M.A.Souza Gustavo, E.A.P.A.Batista, “Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model”, Information Sciences, V: 484, Pages 302-337, 2019 . https://doi.org/10.1016/j.ins.2019.01.076
20. Han Shihand Suchithra Rajendran, “Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation’s Blood Supply”, Hindawy, Volume 2019 | Article ID 6123745 https://doi.org/10.1155/2019/6123745
21. Jason Brownlee, “Deep Learning for Time Series Forecasting, Predict the Future with MLPs, CNNs, and LSTMs in Python”, Machine Learning Mastery, Google Books, 2018. https://machinelearningmastery.com/deep-learning-for-time-series-forecasting/
22. Amal Mahmud, Ammar Mohammed, “A Survey on Deep Learning for Time-Series Forecasting”, In book: Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges, 2021 DOI: 10.1007/978-3-030-59338-4_19
23. José F. Torres, Dalil Hadjout, Abderrazak Sebaa, Francisco Martínez-Álvarez, , and Alicia Troncoso, “Deep Learning for Time Series Forecasting: A Survey”, Big DataVol. 9, No. 1, 2021 https://doi.org/10.1089/big.2020.0159
24. Che, Z., Purushotham, S., Cho, K. et al. Recurrent Neural Networks for Multivariate Time Series with Missing Values. Springer Nature, Sci Rep Article No. 6085, 2018. https://doi.org/10.1038/s41598-018-24271-9

25. Tavakoli, N., Siami-Namini, S., Adl Khanghah, M. et al. An autoencoder-based deep learning approach for clustering time series data. SN Appl. Sci. 2, 937 (2020). https://doi.org/10.1007/s42452-020-2584-8
26. QiyaoWang, AhmedFarahat, ChetanGupta, ShuaiZheng, “Deep time series models for scarce data”, Neurocomputing, Volume 456, Pages 504-518, 2021. https://doi.org/10.1016/j.neucom.2020.12.132
27. G. Spadon, et al.,"Pay Attention to Evolution: Time Series Forecasting With Deep Graph-Evolution Learning" in IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 44, no. 09, pp. 5368-5384, 2022.DOI: 10.1109/TPAMI.2021.3076155
28. Vishal Goar Manoj Kuri Rajesh Kumar Tomonobu Senjyu Editors, “Advances in Information Communication Technology and Computing”, Lecture Notes in Networks and Systems, Proceedings of AICTC Series 2021, Volume 392, Springer, 2021.https://doi.org/10.1007/978-981-19-0619-0
29. Taha Belkhouja, Yan Yan, Janardhan Rao Doppa, “Training Robust Deep Models for Time-Series Domain: Novel Algorithms and Theoretical Analysis”, The Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22), PP.6055-6063, 2022.
30. Ismail Fawaz, H., Forestier, G., Weber, J. et al,” Deep learning for time series classification: a review”, Data Mining & Knowledge Discovery 33, 917–963 (2019). https://doi.org/10.1007/s10618-019-00619-1
31. Wibawa, A.P., Utama, A.B.P., Elmunsyah, H. et al., “Time-series analysis with smoothed Convolutional Neural Network”, J Big Data 9, 44 (2022). https://doi.org/10.1186/s40537-022-00599-y
32. Yi Zheng, Qi Liu, Enhong Chen, Yong Ge, and J. Leon Zhao, “Time series classification using multi-channels deep convolutional neural networks”, c Springer, WAIM 2014, LNCS 8485, pp. 298–310, 2014.134 . https://www.researchgate.net/publication/303250835
33. Guanghua Ren, Yuting Cao, Shiping Wen, higang Zeng, “ A Modified Elman Neural Network with a New Learning Rate Scheme”, Neurocomputing 286, April 2018.
34. DOI: 10.1016/j.neucom.2018.01.046
35. Harold J. Steudel and S. M. Wu, “A Time Series Approach to Queueing Systems with Applications for Modeling Job-Shop In-Process Inventories”, INFORMS, Management Science, Vol. 23, No. 7 , pp. 745-755, 1977.
36. Sven Knoth & Wolfgang Schmid , :Control Charts for Time Series: A Review”, Frontiers in Statistical Quality Control, pp 210–236, (FSQC,volume 7). DOI https://doi.org/10.1007/978-3-7908-2674-6_14
37. Hailin Li, “On-line and dynamic time warping for time series data mining”, International Journal of Machine Learning and Cybernetics 6(1), 2014. DOI: 10.1007/s13042-014-0254-0
38. Allan Tucker, Xiaohui Liu, “A Bayesian network approach to explaining time series with changing structure”, Intelligent Data Analysis 8(5):469-480, 2014. DOI: 10.3233/IDA-2004-8504
39. Zonghua Liu, "Chaotic Time Series Analysis", Mathematical Problems in Engineering, vol. 2010, Article ID 720190, 31 pages, 2010. https://doi.org/10.1155/2010/720190
40. Mohamed M. Ahmed, Ali Ghasemzadeh, Britton Hammit, Nasim Khan, Anik Das, Elhashemi Ali, Rhonda K. Young, Hesham Eldeeb, “Driver Performance And Behavior In Adverse Weather Conditions: An Investigation Using The SHRP2 Naturalistic Driving, Phase 2”, Final Report, Department of Transportation (Grant No: RS07216), Department of Civil and Architectural Engineering, University of Wyoming, USA, 2018.
41. Haque, M.M., Sarker, S. & Dewan, M.A.A. “Driving maneuver classification from time series data: a rule based machine learning approach”, Appl Intell (2022).

https://doi.org/10.1007/s10489-022-03328-3

42. M. Siami, M. Naderpour and J. Lu, "A Mobile Telematics Pattern Recognition Framework for Driving Behavior Extraction", in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 3, pp. 1459-1472, March 2021. DOI: 10.1109/TITS.2020.2971214.
43. Alizadeh, Morteza, "Advanced Data Analytics Methodologies for Anomaly Detection in Multivariate Time Series Vehicle Operating Data", Theses and Dissertations. 5255 (2021). https://scholarsjunction.msstate.edu/td/5255
44. Ma Y, Xie Z, Chen S, Wu Y, Qiao F. Real-Time Driving Behavior Identification Based on Multi-Source Data Fusion. Int J Environ Res Public Health. 2021 Dec 29;19(1):348. DOI: 10.3390/ijerph19010348.
45. Jun Zhang, ZhongCheng Wu,Fang Li, Chengjun Xie, Tingting Ren,Jie Chen, and Liu Liu, "A Deep Learning Framework for Driving Behavior Identification on In-Vehicle CAN-BUS Sensor Data", Sensors (Basel), 19(6): 1356, 2019. DOI: 10.3390/s19061356
46. Hao Li, Junyan Han, Shangqing Li, Hanqing Wang, Hui Xiang, Xiaoyuan Wang, "Abnormal Driving Behavior Recognition Method Based on Smart Phone Sensor and CNN-LSTM", International Journal of Science and Engineering Applications, Volume 11-Issue 01, 2022. DOI: 10.7753/IJSEA1101.1001 www.ijsea.com
47. Kenny Schlegel, Florian Mirus, Peer Neubert, and Peter Protzel, "Multivariate Time Series Analysis for Driving Style Classification using Neural Networks and Hyperdimensional Computing", IEEE Intelligent Vehicles Symposium (IV), Nagoya, Japan, 2021. DOI: 10.1109/IV48863.2021.9576028
48. K.A.De Zoysa Karunathilaka, P.U.Jayaweera, K.G. Kalani, E R.W.P.M. Meththananda, "Driver Behavior Analysis using Vehicular Data", Project Report, University Of Moratuwa Department of Computer Science and Engineering, 2017. http://dilum.bandara.lk/wp-content/uploads/2017/04/Final-Report-OBD2-and-UBI.pdf
49. Driving Analysis. INVERS, invers.com/en/solutions/driving-analysis. Accessed 25 Oct. 2022
50. A. Gregoriades, C. Florides, M. Pampaka, "Driver Behaviour Analysis through Simulation", IEEE International Conference on Systems, Man, and Cybernetics, 13 October 2013. DOI:10.1109/SMC.2013.627
51. D. Davenne, R. Lericollais, P. Philip, "Reliability of simulator driving tool for evaluation of sleepiness, fatigue and driving performance", Accident; analysis and prevention, volume 45, pages 677-82, 2012. DOI:10.1016/j.aap.2011.09.046
52. K. Zinebi, N. Souissi, K. Tikito, Driver Behaviour Analysis Methods: Applications Oriented Study, Proceedings of the 3rd International Conference on Big Data, Cloud and Applications – BDCA, Morocco, Springer, Applied Intelligence, 2022. DOI: 10.1007/s10489-022-03328-3
53. Mario Luca Bernardi, Marta Cimitile, Fabio Martinelli, Francesco Mercaldo, "Driver and Path Detection through Time-Series Classification", Journal of Advanced Transportation, vol. 2018, Article ID 1758731, 20 pages, 2018. https://doi.org/10.1155/2018/1758731
54. Meenakshi Malik, Rainu Nandal, Surjeet Dalal, Vivek Jalglan, and Dac-Nhuong Le, "Deriving Driver Behavioral Pattern Analysis and Performance Using Neural Network Approaches", Intelligent Automation & Soft Computing (IASC), vol.32, no.1, 2022. DOI:10.32604/iasc.2022.02024
55. K. Saleh, M. Hossny and S. Nahavandi, "Driving behavior classification based on sensor data fusion using LSTM recurrent neural networks," 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017, pp. 1-6, DOI:10.1109/ITSC.2017.831783

56. Y. Moukafih, H. Hafidi and M. Ghogho, "Aggressive Driving Detection Using Deep Learning-based Time Series Classification," 2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), 2019, pp. 1-5, DOI: 10.1109/INISTA.2019.8778416.
57. M. E. Mumcuoglu et al., "Driving Behavior Classification Using Long Short Term Memory Networks," 2019 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE), 2019, pp. 1-6, DOI: 10.23919/EETA.2019.8804534.
58. A. Khodairy and G. Abosamra, "Driving Behavior Classification Based on Oversampled Signals of Smartphone Embedded Sensors Using an Optimized Stacked-LSTM Neural Networks," in IEEE Access, vol. 9, pp. 4957-4972, 2021, DOI: 10.1109/ACCESS.2020.3048915.
59. Shokoufeh Monjezi Kouchak; Ashraf Gaffar, "Detecting Driver Behavior Using Stacked Long Short Term Memory Network With Attention Layer", IEEE Transactions on Intelligent Transportation Systems, Volume: 22, Issue: 6, Page(s): 3420 – 3429, 2021. DOI: 10.1109/TITS.2020.2986697
60. Yunyun Zhao, Hongwei Jia, Loafing,Yanjun Qin,Yueyue Wang, "An abnormal driving behavior recognition algorithm based on the temporal convolutional network and soft thresholding", IET Intelligent Transport Systems, Volume37, Issue9, Pages 6244-6261, September 2022. https://doi.org/10.1002/int.22842
61. Jake Grigsby, Zhe Wang, Yanjun Qi, "Long-Range Transformers for Dynamic Spatiotemporal Forecasting" arXiv:2109.12218v2 [cs.LG] 19 May 2022.
62. Jaekyu Lee, Sangyub Lee, Hyosub Choi and Hyeonjoong Cho, "Time-Series Data and Analysis Software of Connected Vehicles", Computers, Materials & Continua, 2021. DOI:10.32604/cmc.2021.015174
63. Jessica Bhalerao , Archana Kadam , Amruta Shinde , Vaishnavi Mugalikar, Himanshu Bhan,"Proposed Design on Driver Behavioral Analysis", International Journal Of Engineering Research & Technology (Ijert) Volume 09, Issue 05, 2020.https://www.ijert.org/proposed-design-on-driver-behavioral-analysis
64. Qiang Meng, Jinxian Weng, "Classification and Regression Tree Approach for Predicting Drivers' Merging Behavior in Short-Term Work Zone Merging Areas", Journal of Transportation Engineering 138(8):1062-1070, August 2012.DOI: 10.1061/(ASCE)TE.1943-5436.0000412
65. Nengchao Lyu, Yugang Wang, Chaozhong Wu, Lingfeng, Peng, Alieu Freddie Thomas "Using naturalistic driving data to identify driving style based on longitudinal driving operation conditions", Journal of Intelligent and Connected Vehicles 5/1 (2022) 17–35 DOI 10.1108/JICV-07-2021-0008
66. Zhengping Li,Kai Zhang,Bokui Chen,Yuhan Dong,Lin Zhang, "Driver identification in intelligent vehicle systems using machine learning algorithms", IET Intelligent Transport Systems, 2018 https://doi.org/10.1049/iet-its.2017.0254
67. Yang Xing, Chen Lv, Zhaozhong Zhang, Huaji Wang, Xiaoxiang Na, Dongpu Cao, Efstathios Velenis, Fei-Yue Wang, "Identification and Analysis of Driver Postures for In-Vehicle Driving Activities and Secondary Tasks Recognition", IEEE Transactions on Computational Social Systems, Volume 5, Issue 1, pp95-108, 2018. DOI: 10.1109/TCSS.2017.2766884
68. Friedrich Lindow, Christian Kaiser, Alexey Kashevnik, Alexander Stocker, "AI-Based Driving Data Analysis for Behavior Recognition in Vehicle Cabin", Proceeding Of The 27th Conference Of Fruct Association, 2021. https://fruct.org/publications/fruct27/files/Lin.pdf
69. Zaky, A.B., Khamis, M.A., Gomaa, W. (2022). "Markov Switching Model for Driver Behavior Prediction: Use Cases on Smartphones". In: Al-Emran, M., Shaalan, K. (eds) Recent Innovations

in Artificial Intelligence and Smart Applications. Studies in Computational Intelligence, vol 1061. Springer, Cham. https://doi.org/10.1007/978-3-031-14748-7_15

70. Wataru Takano, Akihiro Matsushita, Keijiro Iwao, Yoshihiko Nakamura, “Recognition of Human Driving Behaviors based on Stochastic Symbolization of Time Series Signal”, International Conference on Intelligent Robots and Systems IROS 2008, IEEE/RSJ, IEEEXplore, 2008. DOI:10.1109/IROS.2008.4650671
71. M. L. Bernardi, M. Cimitile, F. Martinelli and F. Mercaldo, "Driver Identification: a Time Series Classification Approach," 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-7, 2018. DOI: 10.1109/IJCNN.2018.8489426.
72. Lyu, W, Gonçalves, R, Guo, F et al., “Applying Entropy to Understand Drivers’ Uncertainty during Car-following”, 12th International Conference on Methods and Techniques in Behavioral Research and 6th Seminar on Behavioral Methods, Volume 1, Krakow, Poland, 2021. https://doi.org/10.6084/m9.figshare.13013717.v2
73. Sasan Jafarnejad, “Machine Learning-Based Methods For Driver Identification And Behavior Assessment: Applications For Can And Floating Car Data”, Phd, Docteur De L’université Du Luxembourg En Informatique, 2020. https://orbilu.uni.lu/bitstream/10993/42721/1/PhD_Thesis_Sasan_JAFARNEJAD.pdf
74. Xiaoliang Ma and Ingmar Andréasson, “Behavior Measurement, Analysis, and Regime Classification in Car Following”, IEEE Transactions On Intelligent Transportation Systems, Vol. 8, No. 1, 2007. http://www.diva-portal.org/smash/get/diva2:11471/FULLTEXT01.pdf
75. Ma, X., & Andréasson, I. (2007). Statistical Analysis of Driver Behavior Data in Different Regimes of the Car-Following Stage. Transportation Research Record, , 87–96, 2018(1) https://doi.org/10.3141/2018-12
76. Umberto Fugiglando, Emanuele Massaro, Paolo Santi, Sebastiano Milardo, Kacem Abida, Rainer Stahlmann, Florian Netter, and Carlo Ratti, “Driving Behavior Analysis through CAN Bus Data in an Uncontrolled Environment”, IEEE Transactions On Intelligent Transportation Systems, Vol. 20, NO. 2, 2019.
77. https://senseable.mit.edu/papers/pdf/20181105_Fugiglando-etal_DrivingBehavior_IEEE-TITS.pdf
78. Navin Ramesh Reddy, “Driving Behaviour Classification: An Eco-driving Approach”, Master Thesis, Pervasive Systems, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 2019. https://essay.utwente.nl/80158/1/FinalVersion_Thesis.pdf
79. Romera, E., Bergasa, L.M., Arroyo, R.: Need data for driver behavior analysis? presenting the public uah-driveset. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). pp. 387–392, IEEE (2016). DOI: 10.48550/arXiv.2205.07870
80. Rangesh, Akshay, and Mohan M. Trivedi. "Handynet: A one-stop solution to detect, segment, localize & analyze driver hands." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1103-1110. 2018. https://doi.org/10.48550/arXiv.1804.07834
81. Jingbo Yang, Ruge Zhao, Meixian Zhu, David Hallac, Jaka Sodnik, Jure Leskovec, “Driver2vec: Driver Identification from Automotive Data”, ”, Computer Science, Machine Learning, Cornel University, arXiv:2102.05234 [cs.LG], 2021. https://doi.org/10.48550/arXiv.2102.05234
82. Szilvia Lestyan, “Privacy of Vehicular Time Series Data”, Ph.D. Dissertation, Budapest University of Technology and Economics, Department of Networked Systems and Services, 2022. https://www.crysys.hu/~lestyan/dissertation.pdf

83. David Chunhu Li, Michael Yu-Ching Lin, And Li-Der Chou, “Macroscopic Big Data Analysis and Prediction of Driving Behavior With an Adaptive Fuzzy Recurrent Neural Network on the Internet of Vehicles”, IEEEAccess, Vol.10, pp. 47881-47895, 2022. DOI: 10.1109/ACCESS.2022.3171247
84. Farid Talebloo, “A Practical Deep Learning Approach to Detect Aggressive Driving Behaviour”, M.Sc. Thesis , University of Calgary, Alberta, 2022.
85. Hailong Liu, Tadahiro Taniguchi, Kazuhito Takenaka, Yusuke Tanaka, Takashi Bando, “Reducing the Negative Effect of Defective Data on Driving Behavior Segmentation via a Deep Sparse Autoencoder”, The 5th IEEE Global Conference on Consumer Electronics (IEEE GCCE2016), 2016. DOI:10.1109/GCCE.2016.7800355
86. Arianna Bichicchi, Rachid Belaroussi, Andrea Simone, Valeria Vignali, Claudio Lantieri, And Xuanpeng Li, “Analysis of Road-User Interaction by Extraction of Driver Behavior Features Using Deep Learning”, IEEEAccess, Volume 8, pp 19638- 19645, 2020. DOI:10.1109/ACCESS.2020.2965940
87. Soma Bandyopadhyay, Anish Datta, Shruti Sachan, Arpan Pal, “Unsupervised Driving Behavior Analysis Using Representation Learning and Exploiting Group-based Training”, AI for transportation workshop, AAAI 2022
88. Guillaume Richard. “Transfer Learning methods for temporal data. Machine Learning [cs.LG]”. Université Paris-Saclay, 2021. English. ffNNT : 2021UPASM037ff. fftel-03550415f
89. Xavier Renard, “Time series representation for classification : a motif-based approach”, Data Structures and Algorithms [cs.DS]. Université Pierre et Marie Curie - Paris VI, 2017. English. ffNNT : 2017PA066593ff. fftel-01922186f
90. Hashim, Fatma A., Mai S. Mabrouk and Walid Al-Atabany, “Review of Different Sequence Motif Finding Algorithms.”Avicenna Journal of Medical Biotechnology 11, 130 – 148, 2019.
91. Abdullah Mueen. “Time series motif discovery: dimensions and applications”, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4(2):152–159, 2014. DOI: 10.1002/widm.1119.
92. Abdullah Mueen, Eamonn Keogh, Qiang Zhu, and Brandon Westover, “Exact Discovery of Time Series Motifs”, Proceedings of the SIAM International Conference on Data Mining, SDM 2009, pp.473 – 484, Sparks, Nevada, USA, 2009. DOI:https://doi.org/10.1137/1.9781611972795.41
93. Sahar Torkamani,Volker Lohweg, “Survey on time series motif discovery”, WIREs Data Mining and Knowledge Discovery, Volume7, Issue2, 2017.https://doi.org/10.1002/widm.1199
94. Yifeng Gao, Jessica Lin, “Efficient Discovery of Time Series Motifs with Large Length Range in Million Scale Time Series”, arXiv:1802.04883v1 [cs.DS] 13 Feb 2018. https://cs.gmu.edu/~jessica/publications/HIME_ICDM17.pdf
95. S. Jain, D. Hallac, R. Sosic, J. Leskovec, MASA: Motif-Aware State Assignment in Noisy Time Series Data, in: arXiv:1809.01819 [cs, stat], Anchorage, Alaska, USA, 2019. http://arxiv.org/abs/1809.01819, arXiv: 1809.01819.
96. Yoshiki Tanaka, Kazuhisa Iwamoto, and Kuniaki Uehara. Discovery of Time-Series Motif from Multi-Dimensional Data Based on MDL Principle. Machine Learning, 58(2):269–300, February 2005. ISSN 1573-0565. DOI: 10.1007/s10994-005-5829-2.
97. Farid Talebloo, Emad A. Mohammed, Behrouz Far, “Deep Learning Approach for Aggressive Driving Behaviour Detection”, arXiv:2111.04794 [cs.LG], 2021. https://doi.org/10.48550/arXiv.2111.04794
98. Duong Tuan Anh & Nguyen Van Nhat, "An efficient implementation of EMD algorithm for motif discovery in time series data," International Journal of Data Mining, Modeling and Management,

Inderscience Enterprises Ltd, vol. 8(2), pages 180-194, 2016. http://www.inderscience.com/link.php?id=77159

99. Silva, M. I., & Henriques, R. (2020). Finding manoeuvre motifs in vehicle telematics. Accident Analysis & Prevention, 138, 105467. https://doi.org/10.1016/j.aap.2020.105467
100. E. Romera, L.M. Bergasa and R. Arroyo, "Need Data for Driving Behavior Analysis? Presenting the Public UAH-DriveSet", IEEE International Conference on Intelligent Transportation Systems (ITSC), pp. 387-392, Rio de Janeiro (Brazil), November 2016. http://www.robesafe.uah.es/personal/eduardo.romera/uah-driveset/
101. Yuksel, Asim Sinan; Atmaca, Şerafettin (2019),"Driving Behavior Dataset", Mendeley Data, V1, DOI: 10.17632/jj3tw8kj6h.1
102. Vasili Ramanishka, Yi-Ting Chen,Teruhisa Misu, and Kate Saenko, "Toward Driving Scene Understanding: A Dataset for Learning Driver Behavior and Causal Reasoning", Conference on Computer Vision and Pattern Recognition, CVPR 2018. https://usa.honda-ri.com/HDD
103. Wawage, Pawan (2022), "Driver Behavior Detection Using Smartphone - Dataset", Mendeley Data, V2, doi: 10.17632/9vr83n7z5j.2. https://data.mendeley.com/datasets/9vr83n7z5j/2
104. Yuksel, Asim Sinan; Atmaca, Şerafettin (2020), "Driving Behavior Dataset", Mendeley Data, V2, DOI: 10.17632/jj3tw8kj6h.2. https://data.mendeley.com/datasets/jj3tw8kj6h/2
105. Martin, Manuel. "The Drive and Act Dataset." The Drive and Act Dataset, 27 May 2021, driveandact.com.
106. Yong Zhang, Junjie Li, Yaohua Guo, August 30, 2018, "Vehicle driving behavior", IEEE Dataport, DOI:https://dx.doi.org/10.21227/qzf7-sj04. https://ieee-dataport.org/documents/vehicle-driving-behavior
107. umi001. "GitHub - Umi001/driverBehaviorDataset." GitHub,
108. github.com/umi001/driverBehaviorDataset. Accessed 28 Sept. 2022.
106. "DBNet." DBNet, www.dbehavior.net. Accessed 28 Sept. 2022.
110. Othman, W.; Kashevnik, A.; Ali, A.; Shilov, N. DriverMVT: In-Cabin Dataset for Driver Monitoring including Video and Vehicle Telemetry Information. Data 2022, 7, 62. https://doi.org/10.3390/data7050062 . https://www.mdpi.com/2306-5729/7/5/62
111. Andreas Geiger, Philip Lenz, and Raquel Urtasun, "Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite", Conference on Computer Vision and Pattern Recognition, Geiger 2012CVPR , 2012.
112. https://www.cvlibs.net/datasets/kitti/eval_odometry.php
113. "HCRL - Driving Dataset." HCRL - Driving Dataset, ocslab.hksecurity.net/Datasets/driving-dataset. Accessed 1 Oct. 2022.
114. "IIT - Istituto Di Informatica E Telematica - Sites/Default/Files/Human-behavior-characterization Pdf." IIT - Istituto Di Informatica E Telematica, www.iit.cnr.it/?page_id=-999. Accessed 2 Oct. 2022.
115. J. Ward, S. Worrall, G. Agamennoni, and E. Nebot, "The Warrigal Dataset: Multi-Vehicle Trajectories and V2V Communications," Intelligent Transportation Systems Magazine, IEEE, vol. 6, iss. 3, pp. 109-117, 2014. doi:10.1109/MITS.2014.2315660
116. Schneegass S, Pfleging B, Broy N, Schmidt A, Heinrich F, "A data set of real-world driving to assess driver workload", In:Proceeding the 5th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI'13), p. 150–7, 2013
117. Leila Esmaeili, Alireza Hashemi Golpayegani, A Novel Method for Discovering Process Based on the Network Analysis Approach in the Context of Social Commerce Systems, J. theor.appl.electron.commer.res. vol.16 no.2 Talca mayo 2021. http://dx.doi.org/10.4067/S0718-18762021000200104

118. J. C. A. M. Buijs, B. F. van Dongen and W. M. P. van der Aalst, "Quality dimensions in process Discovery: The importance of fitness, precision, generalization and simplicity", International Journal of Cooperative Information Systems, vol. 23, no. 01, p. 1440001, 2014.DOI:10.1142/S0218843014400012
119. “Time Series - Wikipedia.” Time Series - Wikipedia, 12 Jan. 2021, en.wikipedia.org/wiki/Time_series.
120. Wang, X.; Mueen, A.; Ding, H.; Trajcevski, G.; Scheuermann, P.; Keogh, E., “Experimental comparison of representation methods and distance measures for time series data”, Data Min. Knowl. Discov. 2013, 26, 275–309.
121. Yahyaoui, H.; Al-Mutairi, A. A feature-based trust sequence classification algorithm. Inf. Sci. 2016, 328, 455–484.
122. Ye, J.; Xiao, C.; Esteves, R.M.; Rong, C. Time Series Similarity Evaluation Based on Spearman’s Correlation Coefficients and Distance Measures. In Proceedings of the International Conference on Cloud Computing and Big Data in Asia, Huangshan, China, June 2015; pp. 319–331. 9
123. Zhao, J.; Itti, L. shapeDTW: Shape Dynamic Time Warping. Pattern Recognition. 2018, 74, 171–184.
124. Reza Haghighi Osgouei, Hojin Lee, Seungmoon Choi, “Comparative evaluation of performance measures for human driving skills”, Intelligent Service Robotics Volume 6, Issue 4, pp 169–180, 2013. https://doi.org/10.1007/s11370-013-0134-6