Institute of Transport Economics
Norwegian Centre for Transport Research

Machine learning advancements for vehicle safety systems

Review of technical foundations and applications

Anders Kielland, Anna Piterskaya, Christian Weber

2069/2024

Title:	Machine learning advancements for vehicle safety systems - Review of technical foundations and applications
Tittel:	Fremskritt innen maskinlæring for kjøretøysikkerhetssystemer - Gjennomgang av tekniske grunnlag og anvendelser
Author:	Anders Kielland, Anna Piterskaya, Christian Weber
Date:	12.2024
TØI Report:	2069/2024
Pages:	25
ISSN Electronic:	2535-5104
ISBN Electronic:	978-82-480-2340-1
Project Number:	302327
Funded by:	Norges forskningsråd
Project:	4857 – CRItical SPeed function
Project Manager:	Anders Kielland
Quality Manager:	Frants Gundersen
Commissioning:	Trude Kvalsvik
Research Area:	Machine Learning and Data Science
Keywords:	Artificial intelligence, Car, Machine learning, Sensor data, Traffic safety

Summary

This report summarizes the integration of Machine Learning (ML) in modern vehicle safety applications. Advances in ML have transformed vehicle safety, shifting from traditional rule-based systems to data-driven, adaptive technologies. These applications include advanced driver assistance, predictive maintenance, real-time traffic management, and autonomous driving. ML encompasses a broad spectrum of methodologies and offers flexibility for various implementations, enabling customization for a wide range of tasks. However, while modern ML approaches easily adapt to diversity in data, they also require substantial amounts of data to perform effectively. Furthermore, ML introduces several challenges, particularly the “black box” problem, which raises ethical and regulatory concerns, as well as issues related to privacy and cybersecurity. Addressing these challenges requires research to improve transparency of models, fairness, and trust in ML-driven safety systems. Importantly, the growing availability of vehicle- and traffic-generated data, enabled by Vehicle-to-Everything communication and smart city infrastructure, further highlights ML’s potential for enhancing vehicle safety.

Kort sammendrag

Rapporten oppsummerer integreringen av maskinlæring (ML) i sikkerhetssystemer til moderne kjøretøy. Den senere tids utvikling innen ML har flyttet fokus fra tradisjonelle regelbaserte systemer til datadrevne, adaptive teknologier i bilers sikkerhetssystemene. Disse systemene inkluderer avansert førerassistanse, prediktivt vedlikehold, sanntids trafikkstyring, monitorering av kjøreradferd og autonom kjøring. ML omfatter et bredt spekter av metoder hvilket gir fleksibilitet for ulike implementeringer, som videre muliggjør mange ulike oppgaver. Moderne ML håndterer seg svært godt store variasjoner i data, men krever også store mengder data for å prestere optimalt. ML introduserer også utfordringer knyttet til “black box” problemet til modellene, som etiske og regulatoriske bekymringer, samt problemer relatert til personvern og cybersikkerhet. Disse utfordringene krever videre forskning på transparens i modellene, og rettferdighet og tillit til ML-drevne sikkerhetssystemer. Den økende produksjonen av kjøretøy- og trafikk-genererte data, fra ”Vehicle-to-Everything” kommunikasjon og smarte-byer infrastruktur, fremhever ytterligere ML sitt potensial i videre forbedring av sikkerhet i kjøretøy.

Preface

Machine learning plays a significant role in modern vehicle safety applications. In this report, we provide an introduction to the field of machine learning, the types of data typically used in vehicle safety applications, and the most prominent approaches and concerns.

This report is part of the project 302327 Finding a CRItical Speed function ahead of a road section for vehicles in motion (CriSp), funded by the Norwegian Research Council. Anders Kielland wrote the Introduction and the two sections on Machine Learning Fundamentals and Approaches in Vehicle Safety. Anna Piterskaya and Christian Weber wrote the Data Sources section, while Anna Piterskaya authored the section on Transparency and Ethical Considerations.

Oslo, December 2024
Institute of Transport Economics

Bjørne Grimsrud
Managing Director

Frants Gundersen
Director of Research

Contents

ENGLISH Summary

Machine learning advancements for vehicle safety systems

Review of technical foundations and applications

TØI Report 2069/2024 • Authors: Anders Kielland, Anna Piterskaya, Christian Weber • Oslo • 25 pages

- Development in the Machine Learning (ML) field has transformed vehicle safety systems.
- ML relies on diverse data sources, including vehicle sensors, traffic flow and infrastructure data, weather conditions, and historical crash reports.
- Main ML applications in vehicle safety include:
  - Accident prediction analysis, leading to improved road condition.
  - Real-time adaptive traffic management, optimizing traffic flow.
  - In-vehicle advanced driver assistance systems that support safe vehicle operation.
  - Driver behaviour and fatigue monitoring, detecting drowsiness and distraction.
  - Identification of mechanical failures before they cause accidents.
- Autonomous vehicle systems, improving perception, planning, and decision-making.
- Challenges include the "black box" problem, bias in ML models, privacy concerns, and regulatory constraints, necessitating explainable AI and stronger cybersecurity measures.

As ML rapidly advances, its integration into modern vehicle safety applications has increased, shifting from traditional rule-based systems to more data-driven and adaptive technologies. With the growing adoption of ML in sensor-based technology, predictive analytics, real-time traffic control, driver behavior monitoring, and autonomous driving, these innovations are significantly enhancing road safety and mobility. Key findings show that ML improves accident prevention, optimizes traffic management, and advances driver assistance technologies.

This report provides an overview of the latest advancements in the use of ML in vehicle safety. It covers the methodological foundations, the integration of applications in vehicle safety, and the ethical and regulatory challenges associated with ML-driven systems.

ML is a data-driven process and often relies on high-quality and compressive datasets to function effectively. Data sources include in-vehicle sensor data from Light Detection and Ranging (LiDAR) technologies, radars, cameras, Global Navigation Satellite Systems (GNSSs), and accelerometers, as well as traffic and infrastructure data from surveillance cameras and smart city platforms. Environmental data are recorded via satellite imagery, weather stations, and internet-of-things sensors. Connected vehicle communication, known as Vehicle-to-Everything (V2X), further enable real-time information sharing between vehicles and infrastructure. Proper security and processing efficiency are essential to fully leverage the data sources in ML capabilities in vehicle safety systems.

ML applications in vehicle safety includes a broad spectrum of approaches and methodology, and are used in predictive modelling, anomaly detection, and adaptive decision-making. For example, real-time proactive accident prediction methods leverage ML to assess crash risks, road conditions, and traffic patterns, allowing dynamic speed limits, traffic signal adjustments, and lane direction changes, thus, optimize for a safer traffic flow. Advanced Driver Assistance Systems (ADAS) use ML for collision avoidance, lane-keeping, adaptive cruise control, and emergency braking. Predictive maintenance and vehicle status monitoring help detect mechanical failures, tire wear, and brake issues before they pose safety hazards. The advancement in Neural nets are fundamental in sensor fusion techniques, and object detection and perception in complex driving environments. Driver behaviour and fatigue monitoring rely on ML-based facial recognition, physiological sensors, and steering pattern analysis to detect distraction or drowsiness. Further, personalized safety recommendations analyse individual driving styles to provide tailored feedback, using gamification to encourage safer habits. Finaly, ML will be essential for autonomous vehicle safety, as self-driving cars must operate in a highly dynamic and complex environments, where ML approaches appear to be the only viable solution.

Despite the advancements, ML-driven vehicle safety faces numerous challenges such as data privacy, cybersecurity risks, lack of transparency, and regulatory barriers. The “black box” nature of ML models necessitates new methodologies for decision verification and trust. Bias in ML outcome can result in unjust evaluations of responsibility. Additionally, regulatory frameworks must evolve to align AI-based safety decisions with ethical principles and robust protection measures.

NORSK

Sammendrag

Fremskritt innen maskinlæring for kjøretøysikkerhetssystemer

Gjennomgang av tekniske grunnlag og anvendelser

TØI rapport 2069/2024 • Forfattere: Anders Kielland, Anna Piterskaya, Christian Weber • Oslo, • 25 sider

- Utviklingen innen maskinlæring (ML) har fundamentalt endret sikkerhetssystemer i moderne kjøretøy igjennom forbedret ulykke-forebyggingsanalyse, tilpasset trafikkhåndtering, førerovervåkning og autonom kjøretøystyring i sikkerhetskritiske situasjoner.
- ML bruker ulike datakilder, inkludert kjøretøysensorer, trafikkflyt- og infrastrukturdata, værdata og historisk ulykkesstatistikk.
- Sentrale ML-applikasjoner i kjøretøys sikkerhetssystemer inkluderer:
    - Ulykkesprediksjoner, som fører til forbedret veiinfrastruktur.
    - Sanntid adaptive trafikkstyring for optimalisert trafikkflyt.
    - Førerassistensesystemer som tilrettelegger for sikker kjøring.
    - Monitorering av førers atferd og tilstand, inkludert døsighet og distraksjon.
    - Deteksjon av mekaniske feil før potensielle ulykker oppstår.
    - Systemer i autonome kjøretøy, som forbedrer persepsjon, planlegging og beslutningstaking.
- Utfordringer inkluderer "svart boks" problemet, bias i ML-modeller, personvernhensyn og regulatoriske begrensninger, som krever forklarbar KI og solide cybersikkerhetstiltak.

Den sener tids økning i bruk av ML i moderne kjøretøys sikkerhetssystemer har før til en endring fra tradisjonelle regelbaserte systemer til mer datadrevne og adaptive teknologier. Med økende implementeringen av ML i sensorbasert teknologi, prediktiv analyse, sanntids-trafikkontroll, overvåking av føreradferd og autonome kjøretøy spiller denne utviklingen en sentral rolle i trafikksikkerhet og mobilitet. Utviklingen har gitt bedre ulykkesforebygging, optimalisert trafikkstyring og en betydelig videreutvikling av førerassistensesystemer.

Denne rapporten gir en oversikt over anvendelse av ML i kjøretøysikkerhet. Den gjennomgår fundamentale prinsipper, forståelse og metoder i ML, integreringen i sikkerhetssystemer og etiske og lovmessige utfordringene knyttet til implementering av ML-drevne systemer.

ML er en datadrevet prosess og er ofte avhengig av store mengder data av høy kvalitet for å fungere optimalt. Datakilder inkluderer sensorer som Light Detection and Ranging (LiDAR), radarer, kameraer, globale navigasjonssatellittsystemer (GNSS) og akselerometre, samt trafikk-målinger, overvåkningskameraer og infrastruktur i smarte byer. Omgivelsesdata registreres via satellittbilder, værstasjoner og sensorer fra «internett-of-things». Kommunikasjon mellom kjøretøy og omgivelsene, kjent som «Vehicle-to-Everything» (V2X), muliggjør sanntidsinformasjonsdeling mellom biler og infrastruktur. Robuste datasikkerhetssystemer og rask dataprosessering er også helt avgjørende for å utnytte MLs fulle potensial ML for sikkerhet i kjøretøy.

Bruk av ML i kjøretøysikkerhet omfatter et bredt spekter av tilnærminger og metoder innen prediktiv modellering, deteksjon av unormaliteter og adaptiv beslutningstaking. For eksempel kan sanntids prediksjonsmetoder brukes til proaktiv å vurdere ulykkesrisiko, veiforhold og trafikkbildet. Dette muliggjør dynamiske regulering av fartsgrenser, trafikksignal og endringer i kjørefeltretning resulterende i optimalisert trafikkflyt med økt sikkerhet. Avanserte førerassistansesystemer (ADAS) bruker ML i kjørefeltholder, adaptiv cruisekontroll og sikkerhetsassistanse. Prediktivt vedlikehold og overvåking av kjøretøystatus oppdager mekaniske feil, dekkslitasje og bremseproblemer før de utgjør en sikkerhetsrisikoer. Den sener tids utvikling i nevrale nettverk er helt sentral i sensorfusjonsteknikker og objektdeteksjon i komplekse kjøremiljøer. Overvåking av bilfører sin atferd og oppmerksomhet er basert på ML-drevet ansiktsgjenkjenning, fysiologiske sensorer og analyse av kjøretøyets oppførsel på veien. Videre analyseres individuelle kjørestiler for å gi personlige tilpassede sikkerhetsanbefalinger, også ved hjelp av gamifisering for å oppmuntre til tryggere vaner. Transport utvikler seg i en retning av autonom kjøring og vil ML være avgjørende for sikkerheten ettersom selvkjørende biler må operere i svært dynamiske og komplekse miljøer, der ML-metoder ser ut til å være den eneste funksjonelle løsningen.

Til tross for fremgang står ML-drevet kjøretøysikkerhet overfor en rekke utfordringer, inkludert personvern, cybersikkerhetsrisikoer, mangel på transparens og regulatoriske hindringer. Det er utfordrende å forstå hva som foregår inne i en ML-modell, noe som krever nye metoder og tilnærminger for beslutningsverifikasjon og tillit til sikkerhetssystemene. Bias i ML-modeller kan føre til feilaktig ansvarsfordeling. Regulatoriske rammeverk må videreutvikles for å tilpasses KI-basert beslutningstaking i samsvar med etiske prinsipper og effektive regulering.

1 Introduction

Over the past several decades, the automotive industry has undergone a remarkable technological transformation. Once dominated by purely mechanical systems, modern vehicles now feature a wide array of electronic components, sensors, and software-driven functions. From early safety systems like seatbelts and airbags to today's sophisticated Advanced Driver Assistance Systems (ADAS) and predictive analytics tools, each innovation has contributed to a substantial reduction in road fatalities and injuries. Yet, despite these advances, traffic accidents remain a leading cause of death and disability worldwide, highlighting the need for ever more effective and proactive safety measures (World Health Organization, 2019).

Machine Learning (ML), a branch of AI focused on enabling computers to learn from data and improve over time without being explicitly programmed, has emerged as a key enabler of the next generation of vehicle safety solutions. ML algorithms excel at analysing large, complex datasets and detecting patterns far beyond the capabilities of traditional rule-based systems. These capabilities are particularly relevant in the dynamic and unpredictable environment of road traffic, where vast streams of information from sensors, cameras, connected infrastructure, and historical records must be processed to ensure timely and accurate decision-making.

In practice, ML techniques are being integrated throughout the mobility ecosystem. Within the vehicle, ML models power systems that detect driver drowsiness or distraction, assist with lane-keeping, and even predict when critical components require maintenance. Beyond the vehicle, ML is instrumental in analysing traffic flows, anticipating accident hotspots, and adapting signal timings to reduce congestion and collision risks. Advanced data communication methods, including Vehicle-to-Everything (V2X) connectivity, further expand the scope of ML's influence, fostering real-time cooperation between cars, infrastructure, and other road users.

This holistic approach to safety encompasses not only vehicles and their occupants but also the broader transportation environment, including vulnerable road users such as pedestrians and cyclists. Although the primary focus of this report centres on vehicle applications, it is important to note that ML-based detection and prediction technologies can also enhance safety for all who share the road. Moreover, as cities become "smarter", the integration of ML-driven safety strategies with intelligent traffic lights, connected signage, and sensor-laden infrastructure holds the promise of systemic improvements that extend beyond any single vehicle, potentially reducing accidents at intersections, improving travel times, and enhancing overall urban mobility.

As the industry moves rapidly toward semi- and fully autonomous driving, the importance of ML in ensuring safety will only grow. Autonomous vehicles rely heavily on ML-based perception and control systems to navigate complex road scenarios reliably. Simultaneously, the widespread deployment of connected vehicles and infrastructure generates massive amounts of traffic data, spurring the development of increasingly sophisticated predictive models that can foresee and mitigate risks before they materialize.

In the sections that follow, we will explore the technical foundations and applications of ML in traffic safety, from predictive modelling and driver behaviour analysis to adaptive traffic management and maintenance prediction. We will also consider the ethical, regulatory, and transparency challenges that accompany these advanced systems. Through this examination, the report aims to offer a comprehensive understanding of how ML is reshaping the landscape of traffic safety, and how it may continue to do so in the years to come.

1.1 Acronyms

ABS	Anti-lock Braking System
ADAS	Advanced Driver-Assistance Systems
AI	Artificial Intelligence
CAN	Controller Area Network
CNN	Convolutional Neural Network
EEG	Electroencephalogram
ESC	Electronic Stability Control
GDPR	General Data Protection Regulation
GNSS	Global Navigation Satellite System
LiDAR	Light Detection and Ranging
LIME	Local Interpretable Model-Agnostic Explanations
ML	Machine Learning
NN	Neural Network
OEM	Original Equipment Manufacturer
SHAP	Shapley Additive explanations
V2X	Vehicle-to-Everything

2 Machine Learning Fundamentals

This section provides a basic overview of ML concepts that will facilitate the understanding of its role in vehicle safety technologies. We explore the building blocks that constitute the ML systems, the relevant types of ML, and their relationship with other fields within AI.

2.1 Definitions and General Understanding

Defining ML can be challenging due to its evolving nature and the variety of perspectives within the field. In 1959, Arthur Samuel who popularized the term ML defined it to be a "*field of study that gives computers the ability to learn without being explicitly programmed*"[1]. A widely cited formal definition by Tom M. Mitchell (1997) states:

*"A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P** if its performance at tasks in **T**, as measured by **P**, improves with experience **E**."*

This definition encapsulates the core idea that ML systems use data (experience **E**) to improve their performance on specific tasks (**T**) over time, as evaluated by a performance measure (**P**). For example, suppose we feed an ML algorithm with a large amount of historical car accident data to enhance vehicle safety features. In this scenario:

- **T** is the task of predicting and preventing car accidents.
- **E** is the process of analysing extensive datasets of driving patterns and accident reports.
- **P** is the accuracy with which the algorithm predicts potential safety hazards or collision risks.

As the algorithm processes more data, its ability to predict and help prevent accidents improves, demonstrating the learning process.

In practical use, ML can be viewed as a collection of algorithms and techniques that allow computers to learn from data and improve their performance on specific tasks without being explicitly programmed for those tasks. This latter is an important characteristic that contrasts with expert systems, as were widely used in early AI-driven car safety tools (discussed below).

ML is both rooted in computer science and extensively builds upon established mathematical and statistical methodologies and theories. Additionally, disciplines such as biology, physics, and psychology have significantly contributed to the development of ML.

However, ML is not just new ways of combining established knowledge. ML has introduced fundamentally new theories and algorithms that have advanced the field of AI. Notable examples include Neural Networks (NN) and Reinforcement Learning, both of which are particularly relevant in vehicle safety applications. These methods will be discussed below.

[1] Although it is not an original quote, it remains the most commonly used paraphrased version of Samuel's sentence: "Programming computers to learn from experience should eventually eliminate the need for much of this detailed programming effort". Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, *3*(3), 210-229.

2.2 Fundamental Building Blocks

While multiple components are often required for the successful implementation of ML, three fundamental building blocks generally form its main theoretical and applied foundation. These core components are **data**, **model**, and **optimization**. Additional components include **features and feature engineering**, **evaluation**, and **deployment**. We briefly outline them below.

Data is a key prerequisite for any ML process. It provides the information needed to identify patterns, train models, and make predictions. Data can be structured (e.g., numerical data in spreadsheets) or unstructured (e.g., text, image, audio). The effectiveness of an ML model strongly depends on the quality and relevance of the data, although some models can tolerate a certain level of imperfections.

A **model** is a mathematical representation, often expressed as a function, that learns patterns and relationships present in data. This learning enables the model to make predictions, identify trends, or uncover hidden structures within the data. Models are built upon theoretical constructs derived from statistics and mathematics.

In ML, selecting a model involves choosing the “best model” from a set of candidates, the model space. This space consists of different model types (e.g., linear, non-linear) and the model parameters. However, the model space is not a fixed set of hypotheses; models learn and adapt through an optimization process (described below) by adjusting their parameters. The choice of model space depends on the problem type, the nature of the data, and the desired outcome.

Features are quantifiable properties or characteristics of data that can be used to train a model. **Feature engineering** involves selecting relevant features, transforming raw data into suitable formats, and generating new features from original features to improve model performance. In car safety applications, features might include vehicle speed, distance to nearby objects, driver reaction times, road conditions and traffic density. Effective feature engineering can significantly enhance a model's effectiveness by focusing on factors that most directly impact safe driving.

Optimization is the process of adjusting a model's parameters to achieve a desired objective, such as minimizing prediction error or maximizing the accuracy of traffic safety hazard detection. It plays a crucial role in the learning process by iteratively improving the model’s performance based on feedback from the data. In ML, optimization is often referred to as **training the model**.

Evaluation assesses a model's performance to ensure it generalizes well to data beyond the training set. Best practices involve using multiple well-adapted performance metrics, such as accuracy, precision, recall, F1-score, and mean squared error, to provide a comprehensive and targeted assessment. In safety-critical systems like car safety, rigorous evaluation is especially vital. Poor model performance in such applications can lead to severe consequences, making thorough and careful evaluation an essential step in the development process.

Deployment involves implementing a model in a real-world environment and ensuring its performance remains effective over time. This process typically includes integrating the model into operational systems, such as vehicle safety systems or traffic management infrastructure, where the real-time decision-making takes place. For example, in car safety applications, the model must process data quickly to perform tasks like activating emergency braking. Continuous monitoring is essential to track performance and detect potential issues, such as model drift, which is typically caused by changes in driving patterns or environmental conditions that reduce the model's accuracy. To address such challenges, models are often updated through regular retraining with new data, ensuring they maintain or improve their performance over time.

Together these elements form the core framework for building and training ML systems. They can be summarized as follows:

- **Data**: the available information (e.g., vehicle sensor data, traffic information).
- **Features**: the representation of the data (e.g., speed, distance to obstacles).
- **Model:** The mathematical representation used to quantitatively capture patterns and structures within the data (e.g., regression models, clustering methods).
- **Optimization**: the process of adjusting the model to best fit the data (e.g., training the model to minimize prediction errors).
- **Evaluation**: the measurement of the model's performance (e.g., using accuracy metrics).
- **Deployment**: the practical application of the model (e.g., integrating into vehicle safety systems).

2.3 Algorithms

An algorithm can be considered a step-by-step procedure, or a set of rules used to solve a problem or perform a task. In ML, an algorithm combines a model with an optimization method to learn the model’s parameters from data. Developing and improving algorithms is one of the most actively researched areas in the field of ML.

For instance, a linear regression model specifies a linear relationship between input features and the output variable. The algorithm then applies optimization, such as ordinary least squares or gradient descent, to find the optimal coefficients (i.e., parameters in linear regression) that best fit this linear function to the given data.

A core part of optimization is **regularization**, which involves applying constraints on the model's parameters to prevent overfitting, a situation where the model performs well on training data but poorly on new, unseen data. This is of great importance in car safety applications, where models must generalize well to ensure reliability in diverse driving conditions.

2.4 Models

ML models can be broadly divided into three main categories, often called learning paradigms or learning methods: **supervised learning**, **unsupervised learning**, and **reinforcement learning**, each tailored to different problem settings and data conditions.

Supervised Learning include the models that are trained on labelled data, where each input is paired with a known output. The goal is to establish a mapping between input and output that enables the model to accurately predict output for new unseen input data. This paradigm is relevant in vehicle contexts such as detecting surrounding objects, classifying road structures, recognizing traffic signs, and forecasting traffic float. Classical supervised learning tasks include **classification** (e.g., distin-guishing whether a detected object is a car or a bicycle) and **regression** (e.g., predicting stopping distances).

A wide range of algorithms fall under the category of supervised learning. Among the most classical models are **Linear** and **Logistic Regression**. These are the basic models that, while less commonly used in vehicle safety today, serve as the foundation for more advanced models and can serve as an integral part of other models. Tree models as **Decision Trees and Random Forests**, where the latter is combination of multiple decision threes, handle complex, non-linear patterns, while **Support Vector Machines** find boundaries hyperplane between classes in high-dimensional spaces.

A particularly important group of models in recent years is **ensemble methods**, which combine multiple models to enhance predictive performance and reduce variance. By leveraging diverse models, ensemble methods can capture a wider range of patterns in the data. Three widely used ensemble techniques are **bagging, stacking,** and **boosting**, with **boosting** standing out as the most effective in many predictive tasks. It has won numerous ML competitions and has been extensively implemented in vehicle safety applications.

Neural Networks (NNs) are probably the group of models that have gained most attention in recent years due to their ability to model highly complex relationships in data. They are fundamentally built upon successive layers, each composed of interconnected computational units. This multi-layered architecture enables the network to learn hierarchical representations, where each layer extracts different levels of abstract features from the input data, allowing it to capture intricate patterns and dependencies. Among the most successful NN architectures are **Convolutional NN (CNNs)** and **Recurrent NN**. CNNs excel at image-related tasks, such as object detection, lane detection, and interpreting camera footage in vehicle driving systems. **Recurrent NNs**, on the other hand, are specialized for modelling sequential data, such as sensor time-series or natural language.

Recently, **Transformer models** have revolutionized ML, particularly in natural language processing. They have been highly successful in applications such as large language models, including ChatGPT, developed by OpenAI. The Transformer architecture was introduced in the seminal 2017 paper "Attention Is All You Need" (Vaswani et al., 2017). Unlike traditional sequential models, Transformers leverage self-attention mechanisms, allowing them to dynamically establish relationships between different parts of an input sequence, capturing long-range dependencies and global context. Although initially designed for natural language models, Transformers have also demonstrated significant performance in computer vision tasks. The introduction of **Vision** Transformers has enabled transformer-based models to replace **Convolutional NN** in several visual perception applications. Unlike Convolutional NNs, which process input data using a sliding filter over the image's features, **Vision Transformers** treat images as sequences of patches, similar to how words are processed in language models. This patch-based approach allows transformers to model complex spatial relationships within an image more effectively. Moreover, Vision Transformers are particularly effective in integrating both spatial and temporal data, making them highly valuable for real-time 3D scene reconstruction in vehicle perception systems. Their ability to segment the vehicle's surroundings into distinct objects and characteristics enhances object recognition, semantic segmentation, and scene understanding. This capability is crucial for autonomous driving and vehicle safety systems, where accurate environmental perception is essential for collision avoidance, lane detection, and navigation in dynamic traffic conditions.

Unsupervised learning, unlike supervised learning, does not rely on labeled outputs, allowing models to freely discover hidden structures, patterns, and relationships within the data. This approach is particularly useful for exploratory analysis, such as clustering driving styles to detect risky behavior, identifying unusual sensor readings, or recognizing atypical vehicle behaviors that may indicate mechanical issues.

Clustering algorithms like **K-Means** and **Hierarchical Clustering,** partition data into meaningful groups, helping to reveal complex patterns that can later be labeled for further analysis. Dimensionality reduction techniques, such as **Principal Component Analysis**, streamline large datasets into more manageable forms while preserving essential information, enabling faster and more efficient model training. Additionally, **Autoencoders**, a class of neural networks specialized for unsupervised tasks, are particularly effective for detecting anomalies by identifying deviations from normal patterns, making them valuable for early fault detection in vehicle sensors and predictive maintenance. They can also compress high-dimensional data into a compact representation, reducing storage and transmission costs while retaining key information for decision-making.

Reinforcement Learning differs from supervised and unsupervised learning by focusing on learning through interaction and sequential decision-making rather than from labelled data or static pattern discovery. In Reinforcement Learning, an agent interacts with an environment by taking actions and receiving rewards or penalties as feedback from the environment. The goal is to develop an optimal policy, i.e., a strategy that maximizes cumulative long-term rewards through trial and error. Reinforcement Learning is inherently dynamic, as learning continuously adapts based on new interactions and environmental feedback.

Two prominent approaches in RL are **value-based** and **policy-based** methods. Value-based methods, such as Q-learning, are effective for discrete decision-making, like determining whether to brake or accelerate. Policy-based methods, on the other hand, are particularly suited for continuous control tasks, such as smoothly adjusting a vehicle's steering angle, enabling them to handle dynamic driving conditions more effectively.

In vehicle applications, Reinforcement Learning is particularly useful for tasks such as lane merging, intersection negotiation, adaptive distance control, and collision avoidance manoeuvres. These behaviours are typically learned through extensive training in simulated environments before being deployed in real-world scenarios to ensure safety and efficiency.

Together, these three learning paradigms provide a powerful framework for advancing vehicle safety. By using supervised models for prediction, unsupervised techniques for anomaly detection, and reinforcement learning for adaptive decision-making, intelligent vehicle systems can be developed to enhance safety across diverse driving conditions.

2.5 Explainability

ML models, particularly those with a large number of parameters, often function as "black boxes", where the internal decision-making processes remain opaque. This lack of transparency can introduce biases, erode trust, and complicate regulatory compliance, especially in safety-critical applications. Explainable AI (called XAI) has emerged to address these challenges by providing insights into how ML models generate predictions.

Explainability is essential for evaluating model performance, ensuring fairness, and fostering transparency in ML-driven decision-making. Understanding the rationale behind each decision builds trust between users and developers while promoting ethical AI.

Various techniques enhance explainability. Feature attribution methods like Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) identify key input features that influence a model's output. Model simplification techniques approximate complex models with interpretable alternatives, while visualization tools (e.g., heatmaps, feature importance plots, and saliency maps) provide intuitive insights into model behaviour.

Ensuring explainability in ML-driven safety applications presents unique challenges due to the trade-off between accuracy and transparency. In domains such as autonomous driving and predictive safety, regulatory compliance necessitates interpretable AI. As XAI evolves, it will be pivotal in bridging the gap between high-performance ML models and the need for interpretability, fairness, and trust in AI-driven vehicle safety systems.

2.6 Machine Learning in comparison to Statistics and Expert Systems

AI includes a wide range of methodologies, with ML, statistics, and expert systems each playing significant roles, both independently and in combination. The landscape of AI has evolved substantially over time, with the prominence of different approaches shifting as technology and resources advanced. Today, ML, especially those leveraging large datasets and advanced computational power, dominate much of what is considered "cutting-edge AI." This does not mean that statistics and expert systems have disappeared; rather, their roles have become more specialized, and their influence on mainstream AI is less direct. Understanding the distinctions and areas of overlap among these fields provides a deeper and more comprehensive understanding of AI; both in its historical development and in its modern applications.

ML and statistics share a common mathematical foundation but differ in goals and applications. Statistics primarily focuses on understanding relationships between variables, estimating parameters, and quantifying uncertainty. It emphasizes interpretability, inference, and hypothesis testing, making it central to fields like social sciences, medicine, and econometrics. In contrast, ML prioritizes prediction and automation, using data-driven models to generalize to unseen data. It excels at handling large-scale, high-dimensional datasets and uncovering complex, nonlinear patterns that traditional statistical models may struggle with. ML often prioritize empirical performance over theoretical guarantees, trading interpretability for flexibility and scalability.

While statistics relies on carefully specified models with a small number of predictors, such as using a linear regression to compare medical treatments, ML may use high-dimensional inputs like genomic data and patient histories to predict outcomes without requiring explicit relationships between variables. Additionally, ML leverages high-performance computing to process massive datasets efficiently, whereas traditional statistics is more often associated with moderate-sized, well-curated data and structured theoretical modelling.

Both disciplines are evolving, with ML increasingly incorporating statistical inference and statistics adopting computational advancements. Understanding their strengths allows researchers to apply the most suitable approach for different analytical challenges.

Expert systems are rule-based AI models that rely on predefined logical rules to make decisions, whereas ML derives patterns from data without explicit programming. While expert systems perform well in stable, well-defined domains, they struggle with complexity and variability. In contrast, ML thrives in dynamic environments, making it particularly valuable for vehicle safety applications and other data-rich fields.

Expert systems, which gained prominence in the 1980s and early 1990s, encode human expertise into "if-then" rules or logical inference engines. They are highly interpretable but require manual updates to accommodate new conditions, making them labour-intensive and difficult to scale. In contrast, ML automates pattern discovery, allowing models to continuously adapt and improve as new data emerges, like evolving traffic patterns, new types of road hazards, or shifts in consumer behaviour.

For instance, in vehicle safety applications, modern cars generate vast amounts of sensor data (e.g., LiDAR, cameras, and sensors). Developing an expert system to handle every possible driving condition would require an immense and impractical rule set. ML models, however, can process these high-dimensional inputs autonomously, identifying subtle cues, such as slight changes in lighting or vehicle motion that might be impossible to manually encode.

Despite their differences, ML, expert systems, and statistical methods are complementary. Statistical learning theory underpins many ML algorithms, while hybrid AI approaches integrate rule-based reasoning with ML techniques; for example, combining knowledge graphs with NN models. These hybrid systems leverage the strengths of both paradigms, leading to more robust, scalable, and interpretable AI solutions.

3 Data Sources

In the context of this report, data relevant to vehicle safety serves multiple purposes, including the development of ML models, continuous monitoring, and optimization of these models during vehicle operation. Additionally, data plays a crucial role in real-time decision-making while the vehicle operates in traffic. The initial phase of model development often involves exploratory data analysis, which leads to the selection of useful data, followed by processing and integration of the data into operational ML models within the vehicle's safety program. In this section, we describe how data is collected by the vehicle's onboard systems and received from external sources. These external sources may include infrastructure elements, such as traffic management systems, and environmental factors, such as weather data.

3.1 In-Vehicle Data Acquisition

Modern vehicles are equipped with numerous sensors and generate vast amounts of data. While the application of ML will primarily occur in-car or on the edge, the development of new ML models generally takes place off-site at the vehicle's original equipment manufacturer (OEM), or their suppliers. While our focus here is on safety-related sensors, many other sensors are also employed to control systems, such as combustion and exhaust management in internal combustion engine vehicles or battery management in electric vehicles.

In the following, we provide a brief overview of typical sensors present in modern vehicles that have the potential to deliver data for ML applications. Most of the information on the sensors mentioned in this section is sourced from the Bosch Automotive Handbook (Bosch, 2018).

Electronic Stability Control (ESC) is a collection of functions designed to ensure stable driving within physical limits, such as keeping the vehicle on the path determined by the steering angle. While ESC is not strictly a data source, it is a system that processes inputs and adjusts braking and steering as needed.

Traction Control System reduces excess wheel spin during acceleration. Data from this system can potentially provide information about the current friction experienced between the tires and the road surface.

The **steering angle sensor and throttle position sensor** measure the requested steering input and torque, respectively. Together with yaw rate and lateral acceleration measurements, these sensors provide essential inputs to the ESC system.

An **accelerometer** (e.g., arrays of single sensors like Bosch's SMI860) measures lateral acceleration, yaw rate, and roll rate. These measurements are crucial for the stability system, allowing it to compare the expected (calculated or modelled) path with actual measurements. The expected path is calculated from prior inputs and compared to the new path based on updated inputs. A mismatch exceeding certain thresholds triggers the electronic control unit to counteract the unexpected behaviour and stabilize the vehicle.

Ultrasonic sensors, radar, and LiDAR use echo-based technology to gather information about the vehicle's surroundings across various distance ranges. Ultrasonic sensors are primarily used in parking applications (typically 0.3–2 m), while radar and LiDAR detect obstacles at greater distances (up to 150 m). Radar-based systems use electromagnetic waves in the radio spectrum, while LiDAR uses shorter wavelengths, typically below the visible light spectrum. LiDAR offers the longest ranges but is more susceptible to adverse weather conditions, such as heavy rain or snow. Radar and LiDAR are utilized for longer distances, primarily in adaptive cruise control applications. Radar can also be employed for emergency braking.

Cameras are used to generate computer vision (images and video). They play a key role in lane-keeping assistance and monitoring other road users (e.g., detecting crossing traffic in front of or behind the vehicle and providing lane shift warnings for approaching vehicles). Camera-based systems are crucial for object detection, segmentation, and other visual analyses.

Each wheel is equipped with **a wheel speed sensor**. In vehicles with anti-lock braking system (ABS) and ESC, these sensors provide critical inputs to the ESC system, enabling control of each wheel independently. This functionality helps correct understeering or oversteering, particularly in curves.

The **tachometer** measures the vehicle's overall speed, while **Global Navigation Satellite Systems** (**GNSS) sensors** provide the vehicle's global position.

Additional sensors include **rain sensors** (primarily controlling windshield wipers but also potentially warning of reduced road friction), **external temperature sensors**, and **driver awareness sensors** (e.g., steering input and eye monitoring).

As ML and sensor technologies evolve, the integration of these systems with cloud services and smart infrastructure will continue to advance vehicle safety and efficiency. Nowadays, modern vehicles are already connected to OEM cloud services enabling vast amounts of data to be collected. For instance, Tesla is known to collect data from cameras and other sensors in their vehicles to be used in development of their self-driving program (Harris, 2022). By the end of 2024, with approximately seven million Tesla cars on the road, one can expect extensive coverage of standard driving situations captured by these vehicles. Particularly for heavy goods vehicles, as well as fleets of personal cars (e.g., for home care services), a variety of data types (e.g., GNSS, vehicle speed, load information) can also be collected and stored in cloud services. This data is typically used for logistics purposes, such as route optimization, delivery tracking, fleet performance monitoring, and ensuring timely maintenance.

Within the vehicle, all data communication occurs in accordance with the Controller Area Network (CAN) bus protocol The encoding of these messages is often proprietary to the OEM. However, certain information, such as vehicle speed and specific sensor statuses, is regulated by the OBD-II protocol, which mandates data sharing by law ("SAE J2012 Diagnostic Trouble Code Definitions," 2016).

However, data collection via CAN and OBD for ML purposes is not fully scalable to the volume of data required for effective model development. To address this limitation, various data generation techniques are employed, including data augmentation, simulation, and synthetic data generation. Data augmentation enhances existing data by applying transformations such as rotations, scaling, or noise injection. Simulation involves modelling real-world conditions to create new data that closely resembles actual driving scenarios, leading to synthetic datasets. Additionally, synthetic data can be generated through numerous other approaches, such as generative ML models, which learn patterns from real data and create entirely new samples.

3.2 Vehicle-to-Everything Communication

Nowadays, an increasing number of devices are interconnected through a combination of local networks, direct device-to-device communication and cloud services, enabling vast data exchange. This trend extends to the automobile industry, where technologies like radio signals and satellite navigation are already well-established. Newer technologies enabling two-way communication are also emerging. One such innovation is Vehicle-to-Everything (V2X) communication, a concept that has gained significant traction as technological advancements accelerate.

The core idea behind V2X is the exchange of information between vehicles, infrastructure, and other traffic participants, enabling intelligent decision-making through cloud services (Lv et al., 2024). V2X encompasses subcategories such as Vehicle-to-Vehicle, Vehicle-to-Infrastructure, and Vehicle-to-

Pedestrian communication. These systems establish multi-layered data pipelines that enhance vehicles' environmental awareness and decision-making capabilities.

Many vehicles are already equipped with ADAS, which use multiple sensors to assist drivers with the safe operation of a vehicle. V2X aims to augment existing ADAS solutions with over-the-air messages as additional input data. ML systems can process this data to provide a more comprehensive understanding of a vehicle's surroundings, improving traffic safety by enabling adaptive route planning, cooperative manoeuvring, and real-time hazard alerts. For example, Audi's Traffic Light Information system integrates ML with V2X technology to connect vehicles with traffic signal infrastructure (Glon, 2020). This system provides drivers with real-time traffic information, enabling adaptive route planning, issuing real-time hazard alerts, and facilitating smoother, safer, and more efficient traffic flow.

While ADAS sensors are highly advanced, they have limitation in detecting hazards hidden behind obstacles or beyond their range. V2X communication addresses this limitation by providing real-time data from connected vehicles and infrastructure, including sudden braking events, reported obstacles, adverse weather conditions, and hazards beyond the driver's line of sight. For example, automotive manufacturers like Volkswagen have implemented V2X technology, which enables direct communication between vehicles of the same brand. This allows for early warnings about hazards—such as stalled vehicles or slippery road segments—to be transmitted through the network before the driver's sensors detect them (*Technical milestone in road safety: experts praise Volkswagen's Car2X technology*, 2020). Fixed hazards, like road construction zones or railroad crossings, could also be equipped with transmitters to prevent accidents caused by poor visibility (e.g., fog) or obstructions.

Beyond safety, V2X technology also enhances driver comfort and convenience. It provides information about available parking spaces, optimal routes, and real-time traffic updates. Economically, V2X-enabled ML systems interact with urban objects like traffic lights and roadside sensors, relaying this information to satellite navigation systems to avoid congested areas. These predictive systems optimize signal timings, reduce traffic jams, and facilitate smoother traffic flows, especially in pedestrian-heavy or air-quality-challenged areas.

V2X introduces unprecedented capabilities to the automotive industry, however, several challenges remain. V2X relies on mesh infrastructure, requiring extensive coverage and maintenance, as well as costly transmitters capable of supporting the system. Additionally, V2X-based ML applications face issues such as cybersecurity vulnerabilities, interoperability standards, and the need for scalable, low-latency communication protocols (Kim et al., 2021). Since no universal standard for V2X technology currently exists, designers must address the complexity of creating a safe communication system that performs reliably in high-noise traffic environments. Integrating diverse data sources into a unified system is also a significant challenge, as it demands substantial research, infrastructure, and the safeguarding of sensitive data (Luo, 2020).

As V2X technologies mature and become more widely adopted, their integration with ML models heralds a new era of intelligent, cooperative transportation. By sharing insights to anticipate and mitigate hazards, V2X-driven ML applications promise a safer, more resilient, and efficient traffic ecosystem.

4 Machine Learning Application and Integration in Vehicle Safety

ML has transformed vehicle safety by utilizing vast amounts of data from vehicle sensors, environmental conditions, and road infrastructure. ML plays a role in post-accident analysis, real-time intervention through adaptive traffic management, and risk mitigation during driving. These applications contribute to safer vehicle design, inform traffic safety regulations, and enhance systems for accident prevention and traffic management. Given the breadth of this field, this section highlights the most prevalent methods rather than providing an exhaustive review.

4.1 Proactive Accident Predictions and Adaptive Traffic Management

Proactive accident prevention relies on the ability to predict and identify explanatory variables, patterns and trends associated with an increased likelihood of accidents. By identifying key contributors to crash probability in post-accident data, such as specific roadway designs, adverse weather conditions, or fluctuating traffic volumes, these models can inform the development of targeted interventions and preventative strategies to effectively mitigate accident risks. Furthermore, as transportation networks become more complex and demands on road infrastructure increase, static, one-size-fits-all solutions are often insufficient for maintaining safe and efficient traffic flows. When integrated with real-time environmental monitoring and connected infrastructures, ML enables adaptive traffic management systems that dynamically respond to changing conditions such as weather, lighting, traffic density, and road quality. Such data-driven dynamic interventions can significantly enhance both traffic safety and mobility.

ML models excel in analysing complex spatiotemporal patterns derived from data sources such as combinations of historical crash records, road infrastructure characteristics, and environmental factor. Recent studies have demonstrated the effectiveness of a variety of ML algorithms, from traditional statistical models to advanced NN architectures, in improving accuracy in accident prediction (Ali et al., 2024; Chai et al., 2024). For example, Iranmanesh et al. (2022) identified high crash risk segments in roads using the ensemble models random forest and boosting. The results indicated that boosting is better for identifying crash frequency factors, while random forest exceled at detecting trends and forecasting, with traffic flow rate, road type, and wind speed identified as key influencing variables. A recent review by He et al. (2024) further emphasised the strengths of the recurrent NNs in traffic prediction, showing their effectiveness in capturing temporal dependencies and spatiotemporal patterns in traffic data. Furthermore, date from geographic information system (GIS) in combined with other data types have in ML approaches demonstrated effectiveness in identifying crash hotspots, and in analysing the influence of road attributes such as intersections, speed limits, and curvature on accident likelihood (Ang et al., 2022).

Another upcoming use of ML is monitoring and maintaining the integrity of the road infrastructure itself. ML-based anomaly detection can identify deteriorating pavement conditions, malfunctioning traffic lights, or sensor faults that degrade system performance (Rathee et al., 2023). By proactively initiating maintenance these systems preserve safety and reduce the long-term strain on critical infrastructure. For example, the Norwegian road authority utilized image-based ML for guardrail inspections, significantly reducing the time required compared to previously used manual methods (*Digital Inspection of National and European Roads for the Norwegian Public Roads Administration*, 2025). Additionally, this approach makes the control job a substantially safer operation (*iSi inSight – digital road safety*).

Furthermore, environmental factors, such as weather and lighting conditions, are crucial in improving the accuracy of crash risk predictions. Theofilatos et al. (2019) compared traditional ML methods, such as random forests and decision trees, with NN for real-time crash prediction. Their analysis revealed that variables like rainfall and speed variance significantly influence crash likelihood, with NN achieving a more robust and balanced performance. These findings emphasise the importance of integrating real-time environmental data into proactive safety management strategies.

The incorporation of dynamic, real-time adaptation in traffic management is a field with the potential to revolutionize traffic safety. By integrating large and heterogeneous data streams, ML models can leverage these inputs to forecast near-term traffic conditions, enabling adaptive traffic management systems. For instance, during a sudden downpour that reduces visibility and roadway grip, ML-driven controllers can dynamically adjust speed limits, issue timely public safety warnings, or in the case of changes in road infrastructure, they can divert specific vehicle classes to safer, better-lit corridors or redistribute traffic to less affected routes (Hariharan et al., 2024; Vivekanandan et al., 2024). For example, reinforcement learning is used to optimize adaptive traffic signal control and to provide real-time route recommendations to improve traffic management (Agrahari et al., 2024; Gowri et al., 2024).

In the future, as connected and autonomous vehicles become more prevalent, adaptive traffic management systems may coordinate platoons of driverless cars and dynamically reroute vehicles in response to minor shifts in traffic conditions and safety status. By continuously refining their responses to data sampling, ML-driven systems will become smarter, providing safer and more sustainable mobility. An important and emerging research field is the domain adaptation of ML models, which enables the transfer of knowledge across diverse regions and climates. This advancement addresses the challenge of adapting ML-based traffic management systems to varying environmental and contextual conditions, enhancing their potential in road safety.

4.2 Advanced Driver Assistance Systems

ADAS are technologies which enhance vehicular safety and driving comfort by supporting the human driver in real-time. These technologies include collision warning, collision intervention, driving control assistance, parking assistance, and other driver assistance systems. For example, adaptive cruise control is one of the most popular and relevant applications of ADAS that can benefit from ML (Selvaraj et al., 2023). Further, the systems encompass a variety of functionalities, such as detecting and classifying surrounding objects, maintaining the vehicle’s lane position, and initiating emergency braking manoeuvres, each of which relies on robust perception and decision-making algorithms trained on diverse sensor inputs. ML is emerging as the preferred approach for enabling these functionalities (Ball & Tang, 2019).

CNNs have been pivotal in advancing object detection capabilities in ADAS (Wei et al., 2019). This network type excels in object detection with large scale variation between objects. They enable the integration of contextual information and detailed features, resulting in more accurate detection of objects at various scales, which is necessary for real-time ADAS applications. Pan et al. (2021) developed lightweight CNNs for each of the in-vehicle sensors and then fused these networks using a boosting based strategy. The system leverages redundancies across multiple sensors to provide accurate and reliable object detection. This method not only improves real-time object detection but also maintains high performance even when sensors are partially degraded. By reducing the impact of faulty data and making better use of available sensor inputs, it enhances the safety and reliability during highly diverse driving conditions. Lately, research has explored the application of the transformer NNs (known for their enormous success in large language models) in vehicle image classification within ADAS (Taki & Zemmouri, 2023). By leveraging the self-attention mechanisms inherent in transformers, they have demonstrated competitive performance compared to traditional CNN.

Furthermore, a study applied transformers to particularly difficult vehicle classification problems, revealing an effectiveness in processing low-resolution vehicle images (Dong et al., 2024).

Maintaining correct lane position is crucial for preventing sideswipe collisions and roadway departures. ML-driven lane detection systems often use NNs to conduct a so-called image segmentation. Image segmentation is the process of dividing an image into multiple parts or regions that belong to distinct class categories. This is used to identify lane markings and estimate roadway curvature. For example, Li et al. (2021) developed a lane detection method using CNN, demonstrating robust performance under challenging conditions such as fuzzy or missing lane markings. Furthermore, Mamun et al. (2022) reviewed a set of NN-based frameworks for lane detection, emphasizing the efficacy of CNNs in achieving precise lane detection for ADAS.

When an imminent collision is detected, ML-based ADAS can automatically intervene by e.g. applying emergency braking to mitigate impact severity or prevent crashes entirely. These systems combine predictive modelling with situational awareness, factoring in elements such as object trajectories, relative speeds, and road friction to make informed decisions in real-time. For instance, Itu and Danescu (2023) employed CNNs to predict vehicle velocity and emergency braking events by analysing sequential image data, enabling timely braking responses.

Reinforcement learning has also shown promise, particularly when combined with other ML models. Li et al. (2018) proposed a system in which an NN was used to predict driving track features, while reinforcement learning utilized these features to make steering decisions for the vehicles. The system outperformed the traditional controllers Linear Quadratic Regulators and model predictive control. Additionally, Zhang et al. (2023) introduced a framework combining recurrent NN and reinforcement learning for adaptive control, enabling vehicles to maintain lane position under varying conditions. Furthermore, reinforcement learning have been applied to develop advanced longitudinal control and collision avoidance strategies in high-risk driving scenarios. Chen et al. (2024) demonstrated how these algorithms effectively manage the complexities of ADAS, enhancing the system's ability to navigate hazardous conditions and ensure timely braking.

These studies demonstrated how ML approaches can lead to smoother and more efficient ADAS performance compared to traditional rule-based methods, thereby enhancing overall driving safety.

4.3 Driver Behaviour and Fatigue Monitoring

Monitoring driver attentiveness and detecting signs of fatigue or distraction is increasingly recognized as crucial for accident prevention. According to European Commission (2016) about 95% of road accidents are due to some level of human error, and around 75% are detected to be due to human error alone. In response to regulatory measures, such as the European Union’s mandate on driver monitoring systems for newly produced vehicles (European Commission, 2021a), ML algorithms have become indispensable tools in assessing driver states and delivering timely alerts.

Monitoring driver behaviour and detecting fatigue are critical components of ADAS, aiming to enhance road safety by identifying and mitigating risks associated with driver inattention or drowsiness. NN have been applied to develop systems capable of real-time monitoring and analysis of driver states which in facilitates real-time analysis and response, contributing to the development of proactive in-car alarm systems (Reddy et al., 2023). These systems analyse drivers' physiological and behavioural data, activating alerts when necessary to prevent accidents caused by fatigue or distraction. Furthermore, smartphone-based sensing combined with ML has been explored for driver behaviour classification, providing a cost-effective and accessible means to monitor and improve driving safety (Brahim et al., 2022).

When driver monitoring systems detect signs of fatigue or distraction, they can initiate real-time warnings and interventions to enhance safety. These measures include auditory alerts, haptic feedback like steering wheel vibrations, and visual warnings on the dashboard. In Volvo's driver monitoring system in the EX90 model utilizes cameras to assess driver attentiveness and can autonomously stop the car if impairment is detected (Stokel-Walker, 2024).

Additionally, NN models have been developed for driver behaviour detection, offering robust and accurate predictions that enable timely responses to various driving conditions. For instance, Gheni and Abdul-Rahaim (2024) proposed a model that uses data from in-vehicle sensors. This system captures key driving signals such as acceleration, speed, and throttle position to identify patterns of safe and unsafe behaviour. The model employs a hybrid learning architecture that combines CNNs and recurrent NNs, leveraging the strengths of both models. By analysing real-time data, the approach achieved improved accuracy while maintaining low computational complexity. This innovative method addresses limitations of previous vision-based systems by relying on reliable, privacy-preserving sensor data. ML algorithms are also used to analyse data from in-vehicle cameras, smartphones, and physiological sensors, to classify and predict driver behaviours. By training models on extensive repositories of sensor and vehicle data, these approaches can identify patterns that human analysts might overlook (Abosaq et al., 2022). Further, Divyasri et al. (2024) evaluated a range of boosting algorithms on data collected from in-vehicle sensors to identify driving behaviours as a contributor to road accidents. Sensor features such as engine speed, torque, throttle position, vehicle speed, and steering wheel angle were analysed. The algorithm Light Gradient-Boosting Machine emerged as the most accurate model, achieving an accuracy of 99.68% when classifying between safe and unsafe driving behaviours. By analysing vast datasets from real-world driving scenarios, this research highlights ML’s ability to detect subtle patterns in vehicle dynamics, such as unexpected changes in throttle position or steering wheel inputs, that reflect unsafe behaviour.

Another significant contributor to traffic accidents is a driver fatigue, as it reduces cognitive performance and increases the risk of collisions. Recent advancements in ML have enabled the development of effective, real-time fatigue detection systems. These systems rely on visual and physiological indicators to monitor drivers’ states and issue timely warnings. One approach leverages the most sophisticated CNNs to analyse video data from in-vehicle cameras. Facial features, such as yawning and eye closure, are monitored to identify signs of fatigue. For example, Makhmudov et al. (2024) developed a system that uses a set of classifiers to extract facial regions and advanced image processing algorithms for fatigue detection. Their model achieved an accuracy of 96.54% in identifying drowsiness under varying lighting conditions and facial angles, demonstrating the effectiveness of CNN-based architectures in improving road safety. Wu (2024) presented a vision-based system using facial alignment that identifies distracted and fatigued behaviours using CNNs to measure eye and mouth movements. In addition to visual methods, physiological signals, such as brain activity, are being explored to enhance fatigue detection. For example, systems utilizing electroencephalogram (EEG) data have been developed to detect driver fatigue. EEG measures electrical activity in the brain and can provide direct insights into a driver's cognitive state. However, traditional EEG setups are often impractical for everyday vehicle use due to their complexity and intrusiveness. Recent advancements have led to the development of more user-friendly, wearable EEG devices designed for real-world applications, including driver monitoring (Casson, 2019).

These techniques underline the importance of ML in creating reliable driver monitoring systems, emphasizing their potential to reduce accidents caused by drowsy or distracted driving. As technology evolves, incorporating additional behavioural and physiological data could further improve the accuracy and robustness of fatigue detection systems. The application of ML in monitoring driver behaviour and detecting fatigue plays a pivotal role in enhancing road safety.

4.4 Personalized Safety Recommendations

ML enables the development of personalized safety recommendations tailored to individual driving behaviours and risk profiles. By analysing patterns in a driver's historical performance, including trip data, vehicle telemetry, and reaction times, ML models can identify specific risk factors, such as habitual speeding, abrupt braking, or sharp cornering, and provide targeted feedback to mitigate these risks (Gheni & Abdul-Rahaim, 2024). Personalized recommendations mark a shift from generic safety measures to individualized, data-driven interventions that address unique driver behaviours.

ML models extract features from trajectories, acceleration and braking patterns, steering inputs, and throttle position, to profile drivers based on their behaviour. ML methods, such as clustering and collaborative filtering, are used to group drivers with similar risk profiles and detect outliers who may require tailored interventions. For example, Ferreira et al. (2017) demonstrated the use of smart-phone sensor data to detect risky behaviours like excessive speeding and harsh braking. These insights enable the detection of patterns that generic road safety campaigns might overlook, allowing for more focused and actionable recommendations. Furthermore, clustering (K-mean) algorithms have been utilized to group drivers with similar risk patterns and to pinpoint drivers who may require targeted feedback (Chen et al., 2023).

Once a driver's risk profile is established, adaptive recommender systems generate personalized suggestions, such as optimal speed ranges or adjusted following distances. These recommendations can be delivered through in-car systems, mobile apps, or telematics dashboards. For instance, Zhou et al. (2025) highlighted how integrating real-time contextual data, such as traffic density and weather conditions, into recommender systems ensures that advice remains relevant and dynamic. This approach can warn drivers prone to late braking when approaching known congestion zones or high-risk intersections.

The effectiveness of personalized safety recommendations depends on user engagement and trust. Research shows that drivers are more likely to adopt suggested behavioural changes when the feedback is presented clearly and incrementally. Gamification strategies, such as rewarding safer driving habits, have proven effective in increasing adherence to recommendations (Sumner et al., 2024). Human-in-the-loop approaches, where drivers provide feedback on the system's recommendations, further enhance trust and improve long-term behavioural modification (Zhao et al., 2023).

4.5 Automatic Control of Vehicle Status

Ensuring vehicle reliability is crucial for road safety, as mechanical failures significantly contribute to traffic accidents. Advances in ML have facilitated the development of systems for automatic vehicle status monitoring, enabling proactive detection of maintenance needs and prevention of equipment-related failures.

ML-based predictive maintenance models, such as those using random forests, boosting, or NNs, have been shown to detect deviations from normal operating patterns in vehicles (Prytz, 2014; Zhu et al., 2020). For instance, a study on tire condition utilizing NN to analyse vibration signals obtained through tire rotation under various inflation pressures (Vasan et al., 2023). This approach enabled the instantaneous detection of tire anomalies, facilitating timely maintenance interventions.

Furthermore, the Norwegian company Roadguard, in collaboration with the Norwegian public roads administration, has implemented a ML based system that assesses the condition of tires on passing vehicles. This system employs scanners placed on the road to measure the tread depth of each tire, identifying vehicles with inadequate tires (Elshani, 2024).

4.6 Autonomous Vehicle Safety

The development of fully autonomous vehicles represents a significant shift in traffic safety, evolving from driver-assistance systems to vehicles capable of independently perceiving, deciding, and acting, eliminating the need for human intervention. Central to this advancement is the application of ML, particularly transformer-based models, which have demonstrated exceptional capabilities in supporting autonomous driving systems (Lai-Dang, 2024).

One notable example is the framework introduced by Li et al. (2022), which utilizes transformers to merge spatial and temporal information from multiple cameras. This approach enables the conversion of input data into a comprehensive 3D understanding of the vehicle's surroundings, known as "Bird's Eye View"-perception using a system called BEVFormer. Additionally, this system accurately detects objects and creates segmented maps, even under challenging conditions like low visibility, achieving performance comparable to LiDAR-based systems.

Beyond perception, autonomous driving systems rely on prediction to plan routes, execute lane changes, and execute safe manoeuvres. Reinforcement learning plays a pivotal role in this process, offering the capability to adaptively optimize driving strategies (Booher et al., 2024). In general, the complexity of fully autonomous driving is profound, requiring the vehicle to balance safety, efficiency, and passenger comfort while adapting to dynamic conditions. ML models are expected to play a critical role in managing both reactive and proactive safety measures (Zhu et al., 2020).

While ML-based technology used in human-driven vehicles is also central to safety in autonomous driving, the autonomous context introduces additional layers of complexity and responsibility. Autonomous systems must operate without human intervention, requiring more sophisticated models that can anticipate and respond to a wider range of scenarios.

As of 2024, the industry has made notable progress toward higher levels of vehicle autonomy. For instance, Rivian plans to introduce advanced hands-free driver assistance systems in 2025 and "eyes-off" systems in 2026 to enhance driver convenience and safety (Glovac, 2025). Similarly, Mercedes-Benz's Level 3 "Drive Pilot" allows for hands-free driving under specific conditions but still requires human oversight (*DRIVE PILOT Automated Driving | Mercedes-Benz USA*; Roy & Bajwa, 2025).

However, fully autonomous vehicles (Level 5 autonomy), capable of operating independently under all conditions, are not yet commercially available.

5 Transparency and Ethical Considerations in Machine Learning Safety Tools

Given that road traffic accidents remain one of the leading causes of deaths (World Health Organization, 2019), developing safety tools to mitigate or eliminate these risks has significant societal benefits. However, like any technological advancement, these developments raise important discussions around ethical considerations, transparency, and their broader societal impacts. These concerns affect all stakeholders and stages—from decision-making and implementation to real-world applications and ensuring ethical and legal compliance.

Technologies that rely on ML for decision-making introduce additional challenges and questions: How much trust should we place in these decisions? Are they aligned with our values and societal norms? To what extent should we rely on or override them? Since ML systems are fundamentally data-driven, questions also arise about the origin, balance, and reliability of this data, as well as whether it adequately represents diverse and real-world scenarios (Federal Trade Commission, 2019).

In this section, we will explore key aspects of transparency, regulatory compliance, and ethical, legal, and privacy considerations related to the use of ML safety tools within the automotive sector. This examination aims to highlight how responsible AI deployment can enhance safety while addressing societal concerns and regulatory demands.

5.1 Ethical, Legal, and Privacy Considerations

Ethical concerns are becoming more important as ML models are used in safety-sensitive areas. This is especially true in the automotive industry, where automated decisions can affect public safety. These decisions also influence consumer trust and the reputations of manufacturers and service providers. Ensuring fairness and preventing discriminatory outcomes requires a rigorous examination of training datasets, careful selection of representative samples, and the implementation of fairness-enhancing techniques to minimize biases that could disadvantage certain groups. European Commission has launched Ethics guidelines for trustworthy AI, what determines AI as lawful, ethical and robust (European Commision, 2019).

Privacy protection must also be addressed with utmost care. Organizations must navigate complex legal frameworks, such as the General Data Protection Regulation (GDPR), while safeguarding sensitive data gathered from vehicle sensors, navigation systems, and connected infrastructures. This often involves employing data anonymization, secure communication protocols, privacy-preserving ML methods, and federated learning techniques that reduce the need to centralize potentially identifiable information.

Beyond privacy, safety and reliability concerns are of great importance. Even minor algorithmic errors can have severe consequences on roadways where ML influences the control and navigation of vehicles. This necessitates stringent testing procedures, formal verification methods, and adherence to industry best practices, to ensure that vehicles operate consistently, predictably, and safely under diverse and evolving conditions.

Finally, accountability must be clearly delineated. Establishing frameworks that allocate responsibility when ML systems fail, underperform, or cause harm is essential. This involves maintaining thorough documentation of model development processes, decision-making criteria, and risk assessments, along with having robust incident response procedures.

5.2 Transparency in Decision-Making

Equally important is the need for transparency and explainability. Transparency refers to the openness and clarity with which the workings, decision-making processes, and limitations of ML systems are communicated, while explainability, a subset of transparency, focuses on making the decisions and predictions of ML systems interpretable and understandable to humans. Beyond meeting regulatory and societal expectations, these factors foster confidence that algorithmic decisions—especially those impacting passenger safety and traffic management—are both justifiable and comprehensible. Since many ML systems are considered “black boxes” without clear insight into their proposed solutions, transparency and explainability are critical factors in making these systems understandable to humans, helping to avoid biases and build trust among users and stakeholders.

To address the bias issues and promote responsible and transparent use of ML systems, Mitchell et al. (2019) introduced model cards—a framework for ethically documenting and reporting ML models' performance and intended use. This initiative laid the groundwork for transparency and accountability in ML systems. Building upon this foundation, Kennedy-Mayo and Gord (2024) further expanded the conversation, reclassifying ethical considerations into trustworthiness, risk environments, and risk management, offering a more structured approach to AI safety and ethics.

In order to ensure reliable, trusted outcomes a thorough evaluation of training datasets is required. Moreover, there is a need for a comprehensive documentation of trained ML models which includes metrics that assess bias, fairness, and inclusion. Such documentation as first suggested in the form of “Datasheets for Datasets” by Gebru et al. (2021). This framework emphasizes documenting the motivation, composition, collection process, and recommended use cases of datasets, enabling users to assess their appropriateness and limitations.

Some of the other suggestions to ensure trustworthiness and transparency in ML systems are certification labels proposed by Scharowski et al. (2023). Certification labels are designed to communicate the trustworthiness of ML systems to end users, helping them compensate for their lack of expertise in interpreting the complex material included in the documentation.

Approaches such as explainable AI methodologies (e.g., feature attribution like LIME or SHAP) and research into inherently interpretable model architectures have been gaining popularity during the last years. Techniques under this umbrella aim to make AI systems' decisions interpretable and clarify how inputs influence outputs, thereby making complex models more accessible to stakeholders.

Explainable ML methods have also become of interest in post-accident analysis. Understanding why a model attributes a collision to certain contributing factors (e.g., poor roadway friction versus a distracted driver) is essential for legal investigations, insurance assessments, and policy formulation. Techniques such as SHAP can help interpret the influence of key features, enabling stakeholders to trust the outcomes of automated analyses (Atakishiyev et al., 2024).

5.3 Regulatory Compliance

When it comes to ML-based tools and ML-driven solutions, it is always important to ensure that they comply with regulations and legal standards. Failing to do so can lead to weakened public safety and a loss of trust between end users, regulators, and stakeholders. Nowadays, there are many discussions about regulations for AI (Downes, 2023; Ruggeri, 2024; Stacey, 2023), and many countries are trying to implement guidelines. The EU has recently passed an AI law called the “AI Act” (European Commision, 2019; European Commission, 2021b), while the US-based National Institute of Standards and Technology (NIST) has released “Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile” (National institute of Standard and technology, 2024). There are different aspects of ML-based technology that should be regulated. Here, we focus on the main issues and provide a general overview.

First, safety standards are vital for the transport sector, and it is important to align technology with established standards to eliminate risks and provide reliable solutions. The International Organization for Standardization (ISO) has provided some guidelines, such as ISO 26262-1:2018 (Road vehicles — Functional safety), which describes safety-related systems that include one or more electrical and/or electronic systems installed in series production road vehicles (International Organization for Standardization, 2018), and ISO 21448:2022 (Road vehicles — Safety of the intended functionality), which provides guidance on measures to ensure the Safety of the Intended Functionality (SOTIF), i.e., the absence of unreasonable risk due to a hazard caused by functional insufficiencies (International Organization for Standardization, 2022). Not only should rigorously testing and validation procedures be incorporated to ensure reliable performance under different conditions and failures, but fail-safe mechanisms should also be implemented to prevent undesired outcomes in cases of malfunction.

The European Union has implemented legislation mandating the inclusion of driver monitoring systems in new vehicles to enhance road safety. As of July 2022, all new car models in the EU are required to be equipped with driver drowsiness and attention warning systems designed to monitor driver alertness and provide warnings to prevent accidents caused by fatigue or distraction (European Commission, 2021a).

Automotive manufacturers have responded by integrating machine learning-based solutions into their vehicle designs to comply with these regulations and improve overall safety. These systems utilize advanced algorithms to detect signs of driver drowsiness or distraction, thereby enhancing the vehicle's ability to prevent accidents (Kruszynska, 2024).

While initial implementations focus on detecting drowsiness and distraction, ongoing research is exploring the detection of additional driver states, such as emotional stress or aggression, that may influence driving behaviour. Advancements in AI and ML are enabling the development of more sophisticated driver monitoring systems capable of assessing a wider range of cognitive and emotional states, further contributing to road safety (Qu et al., 2024).

Secondly, ML-based solutions should be explainable and transparent to demonstrate clear accountability. Providing necessary documentation that describes the design of ML algorithms and models, as well as information about training data, is one of the main requirements subject to audits by regulatory bodies. Critical safety systems should produce human-interpretable, explainable outputs. This further implies regulations for clear accountability in cases of failures or accidents, as well as requirements for reporting such cases to regulatory bodies.

Thirdly, as all ML-based solutions rely on big data. Regulations should be in place to ensure that there is no bias or discrimination, enabling transport tools to securely and fairly operate while providing equitable outcomes across different scenarios (Federal Trade Commission, 2019). Data used in training should also comply with data privacy laws, such as the General Data Protection Regulation (GDPR). Such data should not only be anonymized but also encrypted and securely stored to protect individuals' identities and prevent data breaches. Regulations ensuring privacy, security, and fairness must be firmly established for any further AI technology development in the transport sector.

Regulatory compliance is essential for the implementation, deployment, and adoption of ML-based technologies. Ensuring that such systems are fair, comply with ethical requirements, and secure public safety and trustworthiness is vital for the continued advancement of transportation technologies.

Taken together, these factors underscore the necessity of an ethically grounded, well-regulated, transparent, and accountable approach to ML in the automotive sector and beyond. Ultimately, the sustainable adoption of these technologies depends on public trust, safety, and respect for fundamental human values.

6 Conclusion or Future Directions

The automotive sector is undergoing transformative advancements that are redefining how we drive and experience safety on the roads. ML is at the forefront of these innovations, enabling smoother traffic navigation, timely alerts, and proactive risk prevention and mitigation. As ML technologies continue to evolve, their impact is expected to extend beyond individual vehicles to broader transportation systems. Emerging integrations with intelligent traffic signals, connected road signage, and infrastructure sensors—key components of smart city initiatives—promise to enhance safety on a systemic level. Moreover, the safety of vulnerable road users, including pedestrians, cyclists, and motorcyclists, can benefit significantly from ML-driven detection and prediction tools.

Despite the rapid progress in ML applications, several challenges remain before full-scale integration into the automotive sector can be achieved. While ML-based technologies offer promising predictive capabilities, concerns about their robustness persist. Ensuring the reliability of vehicle safety systems will require addressing issues such as cybersecurity, data privacy, and interoperability across manufacturers and other stakeholders. Additionally, fine-tuning ML algorithms demands complex multi-objective optimization and further research into explainable ML and ensemble techniques to improve both accuracy and usability of predictive models. Continuous sensor updates and calibration are also essential to ensure the collection of reliable, actionable data, which can be integrated with other technologies to enhance the dependability and trustworthiness of safety systems.

Looking ahead, ML applications in traffic safety are likely to expand further, seamlessly integrating with smart city infrastructure. Real-time data from connected intersections and adaptive traffic signals could complement ML-driven vehicle systems to create safer and more efficient traffic flows. Advancements in sensor technology and computer vision will also improve the detection and protection of vulnerable road users, ensuring that safety measures extend to all participants in the transportation ecosystem. These developments have the potential to revolutionize traffic safety and make roads safer for everyone.

References

Abosaq, H. A., Ramzan, M., Althobiani, F., Abid, A., Aamir, K. M., Abdushkour, H., Irfan, M., Gommosani, M. E., Ghonaim, S. M., & Shamji, V. (2022). Unusual driver behavior detection in videos using deep learning models. *Sensors*, *23*(1), 311.

Agrahari, A., Dhabu, M. M., Deshpande, P. S., Tiwari, A., Baig, M. A., & Sawarkar, A. D. (2024). Artificial intelligence-based adaptive traffic signal control system: A comprehensive review. *Electronics*, *13*(19), 3875.

Ali, Y., Hussain, F., & Haque, M. M. (2024). Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. *Accident Analysis & Prevention*, *194*, 107378.

Ang, K. L.-M., Seng, J. K. P., Ngharamike, E., & Ijemaru, G. K. (2022). Emerging technologies for smart cities' transportation: geo-information, data analytics and machine learning approaches. *ISPRS International Journal of Geo-Information*, *11*(2), 85.

Atakishiyev, S., Salameh, M., Yao, H., & Goebel, R. (2024). Explainable artificial intelligence for autonomous driving: A comprehensive overview and field guide for future research directions. *IEEE Access*.

Ball, J. E., & Tang, B. (2019). Machine learning and embedded computing in advanced driver assistance systems (ADAS). In (Vol. 8, pp. 748): MDPI.

Booher, J., Rohanimanesh, K., Xu, J., Isenbaev, V., Balakrishna, A., Gupta, I., Liu, W., & Petiushko, A. (2024). CIMRL: Combining IMitation and Reinforcement Learning for Safe Autonomous Driving. *arXiv preprint arXiv:2406.08878*.

Bosch, R. (2018). *Bosch automotive handbook* (10 ed.). SAE International.

Brahim, S. B., Ghazzai, H., Besbes, H., & Massoud, Y. (2022). A machine learning smartphone-based sensing for driver behavior classification. 2022 IEEE International Symposium on Circuits and Systems (ISCAS),

Casson, A. J. (2019). Wearable EEG and beyond. *Biomedical engineering letters*, *9*(1), 53-71.

Chai, A. B. Z., Lau, B. T., Tee, M. K. T., & McCarthy, C. (2024). Enhancing road safety with machine learning: Current advances and future directions in accident prediction using non-visual data. *Engineering Applications of Artificial Intelligence*, *137*, 109086.

Chen, D., Gong, Y., & Yang, X. (2024). Deep Reinforcement Learning for Advanced Longitudinal Control and Collision Avoidance in High-Risk Driving Scenarios. *arXiv preprint arXiv:2404.19087*.

Chen, S., Cheng, K., Yang, J., Zang, X., Luo, Q., & Li, J. (2023). Driving behavior risk measurement and cluster analysis driven by vehicle trajectory data. *Applied Sciences*, *13*(9), 5675.

Digital Inspection of National and European Roads for the Norwegian Public Roads Administration. (2025). https://isi.no/post/2025/01/digital-inspection-of-national-and-european-roads-for-the-norwegian-public-roads-administration/

Divyasri, C., Neelima, N., & Smitha, T. (2024). Machine Learning for Road Safety Enhancement Through In-Vehicle Sensor Analysis. 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT),

Dong, X., Shi, P., Tang, Y., Yang, L., Yang, A., & Liang, T. (2024). Vehicle Classification Algorithm Based on Improved Vision Transformer. *World Electric Vehicle Journal*, *15*(8).

Downes, B. L. a. L. (2023). *Who Is Going to Regulate AI?* https://hbr.org/2023/05/who-is-going-to-regulate-ai

DRIVE PILOT Automated Driving | Mercedes-Benz USA. https://www.mbusa.com/en/owners/manuals/drive-pilot

Elshani, N. (2024). *A system placed on the road evaluates the condition of tires on vehicles - Telegrafi*. https://telegrafi.com/en/a-system-installed-on-the-road-evaluates-the-condition-of-tires-in-vehicles/

European Commision. (2019). *Ethics guidelines for trustworthy AI*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission. (2016). *REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL Saving Lives: Boosting Car Safety in the EU Reporting on the monitoring and assessment of advanced vehicle safety features, their cost effectiveness and feasibility for the review of the regulations on general vehicle safety and on the protection of pedestrians and other vulnerable road users*. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52016DC0787

European Commission. (2021a). *Commission Delegated Regulation (EU) 2021/1341 of 23 April 2021 supplementing Regulation (EU) 2019/2144 of the European Parliament and of the Council by laying down detailed rules concerning the specific test procedures and technical requirements for the type-approval of motor vehicles with regard to their driver drowsiness and attention warning systems and amending Annex II to that Regulation (Text with EEA relevance)*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32021R1341

European Commission. (2021b). *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0206

Federal Trade Commission. (2019). Big Data-A tool for inclusion or exclusion? Understanding the issues (2019).

Ferreira, J., Carvalho, E., Ferreira, B. V., de Souza, C., Suhara, Y., Pentland, A., & Pessin, G. (2017). Driver behavior profiling: An investigation with different smartphone sensors and machine learning. *PLoS one*, *12*(4), e0174959.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, *64*(12), 86-92.

Gheni, H. M., & Abdul-Rahaim, L. A. (2024). An Efficient Deep Learning Model Based on Driver Behaviour Detection Within CAN-BUS Signals. *Revue d'Intelligence Artificielle*, *38*(1).

Glon, R. (2020). *Audi taught cars to talk to traffic lights, and they're set to be even chattier*. https://www.digitaltrends.com/cars/audi-v2x-traffic-light-technology

Glovac, J. (2025). *Rivian Targets Hands-Free Driving in 2025, Eyes-Off Technology by 2026*. https://wheelfront.com/rivian-targets-hands-free-driving-in-2025-eyes-off-technology-by-2026/

Gowri, P., Sivapriya, G., Venkateswaran, K., Sridhar, N., Indhumathi, N., & Sathya, M. (2024). Adaptive Traffic Control Using Machine Learning Algorithm. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT),

Hariharan, R., Balaji, T., Balaji, S., & Bharathi, S. P. (2024). Adaptive Machine Learning Algorithms and Models for Real-Time Prediction of Traffic Flow in Urban Areas. 2024 International Conference on Cybernation and Computation (CYBERCOM), Harris, M. (2022). *The Radical Scope of Tesla's Data Hoard*. https://spectrum.ieee.org/tesla-autopilot-data-scope

He, Y., Huang, P., Hong, W., Luo, Q., Li, L., & Tsui, K.-L. (2024). In-Depth Insights into the Application of Recurrent Neural Networks (RNNs) in Traffic Prediction: A Comprehensive Review. *Algorithms*, *17*(9), 398.

International Organization for Standardization. (2018). *ISO 26262-1:2018 (Road vehicles — Functional safety)*. https://www.iso.org/standard/68383.html

International Organization for Standardization. (2022). *ISO 21448:2022 (Road vehicles — Safety of the intended functionality)*. https://www.iso.org/standard/77490.html

Iranmanesh, M., Seyedabrishami, S., & Moridpour, S. (2022). Identifying high crash risk segments in rural roads using ensemble decision tree-based models. *Scientific Reports*, *12*(1), 20024.

iSi inSight – digital road safety. https://isi.no/en/insight/

Itu, R., & Danescu, R. (2023). Fully convolutional neural network for vehicle speed and emergency-brake prediction. *Sensors*, *24*(1), 212.

Kennedy-Mayo, D., & Gord, J. (2024). " Model Cards for Model Reporting" in 2024: Reclassifying Category of Ethical Considerations in Terms of Trustworthiness and Risk Management. *arXiv preprint arXiv:2403.15394*.

Kim, K., Kim, J. S., Jeong, S., Park, J.-H., & Kim, H. K. (2021). Cybersecurity for autonomous vehicles: Review of attacks and defense. *Computers & security*, *103*, 102150.

Kruszynska, M. (2024). *Driver Monitoring Systems to become mandatory under new EU and US road safety regulations*. https://spyro-soft.com/blog/automotive/driver-monitoring-systems-to-become-mandatory-under-new-eu-and-us-road-safety-regulations

Lai-Dang, Q.-V. (2024). A Survey of Vision Transformers in Autonomous Driving: Current Trends and Future Directions. *arXiv preprint arXiv:2403.07542*.

Li, D., Zhao, D., Zhang, Q., & Chen, Y. (2018). Reinforcement learning and deep learning based lateral control for autonomous driving. *arXiv preprint arXiv:1810.12778*.

Li, J., Zhang, D., Ma, Y., & Liu, Q. (2021). Lane image detection based on convolution neural network multi-task learning. *Electronics*, *10*(19), 2356.

Luo, F.-L. (2020). Machine learning for future wireless communications.

Lv, S., Qin, Y., Gan, W., Xu, Z., & Shi, L. (2024). A systematic literature review of vehicle-to-everything in communication, computation and service scenarios. *International Journal of General Systems*, 1-31.

Makhmudov, F., Turimov, D., Xamidov, M., Nazarov, F., & Cho, Y.-I. (2024). Real-Time Fatigue Detection Algorithms Using Machine Learning for Yawning and Eye State. *Sensors*, *24*(23), 7810.

Mamun, A. A., Ping, E. P., Hossen, J., Tahabilder, A., & Jahan, B. (2022). A comprehensive review on lane marking detection using deep neural networks. *Sensors*, *22*(19), 7682.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. Proceedings of the conference on fairness, accountability, and transparency,

Mitchell, T. M. (1997). *Machine learning* (Vol. 1). McGraw-hill New York.

National institute of Standard and technology. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

Pan, H., Sun, W., Sun, Q., & Gao, H. (2021). Deep learning based data fusion for sensor fault diagnosis and tolerance in autonomous vehicles. *Chinese Journal of Mechanical Engineering*, *34*(1), 72.

Prytz, R. (2014). *Machine learning methods for vehicle predictive maintenance using off-board and on-board data* Halmstad University Press].

Qu, F., Dang, N., Furht, B., & Nojoumian, M. (2024). Comprehensive study of driver behavior monitoring systems using computer vision and machine learning techniques. *Journal of Big Data*, *11*(1), 32.

Rathee, M., Bačić, B., & Doborjeh, M. (2023). Automated road defect and anomaly detection for traffic safety: a systematic review. *Sensors*, *23*(12), 5656.

Reddy, R. M., Konguvel, E., & Sivakumar, R. (2023). A Proactive In-Car Alarm System for Driver Fatigue Detection Using Machine Learning. 2023 International Conference on Next Generation Electronics (NEleX),

Roy, A., & Bajwa, A. (2025). *Rivian to launch hands-free driving system in 2025, 'eyes-off' in 2026, CEO says*. https://www.reuters.com/business/autos-transportation/rivian-launch-hands-free-driving-system-2025-ceo-says-2025-01-23/

Ruggeri, A. (2024). *Davos 2024: Can – and should – leaders aim to regulate AI directly?* https://www.bbc.com/worklife/article/20240118-davos-2024-can-and-should-leaders-aim-to-regulate-ai-directly

SAE J2012 Diagnostic Trouble Code Definitions. (2016). In: SAE International.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, *3*(3), 210-229.

Scharowski, N., Benk, M., Kühne, S. J., Wettstein, L., & Brühlmann, F. (2023). Certification labels for trustworthy ai: Insights from an empirical mixed-method study. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency,

Selvaraj, D. C., Hegde, S., Amati, N., Deflorio, F., & Chiasserini, C. F. (2023). A deep reinforcement learning approach for efficient, safe and comfortable driving. *Applied Sciences*, *13*(9), 5272.

Stacey, D. M. a. K. (2023). *Five takeaways from UK's AI safety summit at Bletchley Park*. https://www.theguardian.com/technology/2023/nov/02/five-takeaways-uk-ai-safety-summit-bletchley-park-rishi-sunak

Stokel-Walker, C. (2024). *Making Safer Drivers: Volvo Cars Driver Understanding System*. https://time.com/7094834/volvo-cars-driver-understanding-system

Sumner, E. S., DeCastro, J., Costa, J., Gopinath, D. E., Kimani, E., Hakimi, S., Morgan, A., Best, A., Nguyen, H., & Brooks, D. J. (2024). Personalizing driver safety interfaces via driver cognitive factors inference. *Scientific Reports*, *14*(1), 18058.

Taki, Y., & Zemmouri, E. (2023). Vehicle Image Classification Method Using Vision Transformer. International Conference on Artificial Intelligence & Industrial Applications,

Technical milestone in road safety: experts praise Volkswagen's Car2X technology. (2020). https://www.volkswagen-newsroom.com/en/press-releases/technical-milestone-in-road-safety-experts-praise-volkswagens-car2x-technology-5914?utm_source=chatgpt.com

Theofilatos, A., Chen, C., & Antoniou, C. (2019). Comparing machine learning and deep learning methods for real-time crash prediction. *Transportation research record*, *2673*(8), 169-178.

Vasan, V., Sridharan, N. V., Prabhakaranpillai Sreelatha, A., & Vaithiyanathan, S. (2023). Tire condition monitoring using transfer learning-based deep neural network approach. *Sensors*, *23*(4), 2177.

Vivekanandan, G., Pathmanaban, J., James, P. G., Priyanka, V., & Samvardhini, V. (2024). Smart Traffic Management System for Urban Mobility Enhancement Using RF Sensors, Cameras, and Machine Learning. 2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS),

Wei, J., He, J., Zhou, Y., Chen, K., Tang, Z., & Xiong, Z. (2019). Enhanced object detection with deep convolutional neural networks for advanced driving assistance. *IEEE transactions on intelligent transportation systems*, *21*(4), 1572-1583.

World Health Organization. (2019). *Global status report on road safety 2018*. World Health Organization.

Wu, D. (2024). Improving automatic detection of driver fatigue and distraction using machine learning. *arXiv preprint arXiv:2401.10213*.

Zhang, Y., Boedecker, J., Li, C., & Zhou, G. (2023). Incorporating Recurrent Reinforcement Learning into Model Predictive Control for Adaptive Control in Autonomous Driving. *arXiv preprint arXiv:2301.13313*.

Zhao, Z., Liao, X., Abdelraouf, A., Han, K., Gupta, R., Barth, M. J., & Wu, G. (2023). Real-time Learning of Driving Gap Preference for Personalized Adaptive Cruise Control. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC),

Zhou, X., Liao, Z., Zhao, Y., Liu, Y., & Yi, A. (2025). Ride-hailing pick-up area recommendation in a vehicle-cloud collaborative environment: a feature-aware personalized clustering federated learning approach. *Cluster Computing*, *28*(1), 32.

Zhu, M., Wang, Y., Pu, Z., Hu, J., Wang, X., & Ke, R. (2020). Safe, efficient, and comfortable velocity control based on reinforcement learning for autonomous driving. *Transportation Research Part C: Emerging Technologies*, *117*, 102662.

The Institute of Transport Economics (TØI) is an applied research institute that carries out research and study assignments for businesses and public agencies. TØI was established in 1964 and is organized as an independent foundation. TØI develops and disseminates knowledge about transport with scientific quality and practical application. The Institute has an interdisciplinary environment with 90+ highly specialized researchers.

The Institute conducts research dissemination through TØI reports, articles in scientific journals, books, seminars, as well as posts and interviews in the media. TØI reports are available free of charge on the Institute's website www.toi.no.

TØI covers all means of transport and thematic areas within transport, including traffic safety, public transport, climate and environment, tourism, travel habits and travel demand, urban planning, ITS, public decision-making processes, business transport and general transport economics. The Institute participates actively in international research collaboration, with particular emphasis on the EU framework programs.

The Institute of Transport Economics requires copyright for its own work and emphasizes acting independently of the clients in all professional analyses and assessments.

Postal Address:
Institute of Transport Economics
P.O. Box 8600 Majorstua
N-0349 Oslo
Norway

Business Address:
Forskningsparken
Gaustadalléen 21

Email: toi@toi.no

Web address: www.toi.no