Ethical Implications of Machine Learning Algorithms for Autonomous Vehicle Decision-Making

By Dr. Chukwuemeka Eneh

Professor of Electrical Engineering, University of Benin, Nigeria

1. Introduction

[1] Since the introduction of the first commercial autonomous vehicles (AVs) in the early 2010s, the development of this technology has shifted towards increasingly higher autonomy levels at a rapid pace. The% current development of highly or fully automated vehicles and Advanced Driver Assistance Systems (ADAS) improves safety and efficiency for drivers and stakeholders, but it also raises numerous ethical and legal questions that need to be addressed along the way to market introduction [2, 3, 4]. The principle of protecting human life as high as possible gained paramount importance throughout the public discussion and vehicle manufactures are in a hurry to constraints decisions so designed cars will do the 'right thing' in critical situation [6, 7, 8].[2] Machine learning (ML) underpins many commercially-available AI systems – including recommendation systems, autonomous vehicles and AI-augmented medical diagnosis – but has been blamed for discriminatory algorithms that favor majority groups, and for AI misuse by malicious actors. This review surveys the primary risks associated with AI systems and recent progress toward their mitigation noting that self-training mechanisms provide an alternative to black-box models for AI augmentation. Beyond these short-term risks, however, there are longer-term speculative concerns that AI could subvert some human interests. Effective regulation cannot rely on technical expertise alone and social scientists must educate the broader public about the nature and possible solutions oriented to these risks.

1.1. Background and Significance

The rapidly growing interest and dominance of autonomous vehicles have raised ethical and moral concerns among society, governments, and other concerned communities. Each day some new accidents are added to the database of accidents caused by autonomous vehicles. These severe and unfortunate accidents have raised new ethical, moral, and legal concerns for the autonomous vehicle researchers, manufacturers, and regulatory authorities. Autonomous systems with superior design may take similar decisions as human do, and those were recognized and adopted by human unconditionally. The algorithm-based decisions in autonomous vehicles are made considering the data of numerous accidents, hence the decisions of autonomous vehicle are more plausible than those made by humans. If an accident or even a severe casualty does occur because decisions are made by autonomous vehicles, the responsible party is the manufacturer or the designer of the autonomous vehicle, the algorithmic designer, manager, and the owner of the vehicle [3].

The number of vehicles on roads every year is increasing swiftly, and unfortunately, the number of collisions and casualties is also increasing in parallel. Many reported accidents happen due to human errors and negligence, and these accidents often cause financial loss as well as severe casualties. To handle this crisis and to reduce the number of accidents and causalities, researchers have developed autonomous efficient algorithm-based systems. Such systems are recognized now as autonomous vehicles, and many renowned researchers have shown a keen interest in involving their research endeavors in the enhancing and betterment of such autonomous systems. An autonomous vehicle provides the only solution for minimizing human mistakes while they are driving vehicles. Such fully autonomous functionality enables the vehicle to make important decisions without human involvement while driving [4].

1.2. Research Objectives

The autonomous vehicle sector has been studied for decades while technological advances have made self-driving a ubiquitous topic, contributing profusely to numerous domains spanning beyond the Computer Science. However, researchers have already spent colossal manpower and time to increase the efficiency and robustness of these vehicles, and more and more critical decisions about the lives of the driver and those nearby are becoming automated. Accordingly, one of the most important challenges facing the autonomous vehicle industry is to design AI that can make ethical decisions about vehicle control. As an inevitable continuation of these design choices, at the end of this ambitious study, we attempt to answer several challenging questions, such as how the ethical axioms presented in the Trolley Dilemma can effectively be integrated without tremendous overlap, how the autonomous vehicle can effectively assess and quantify human susceptibilities, if the autonomous vehicle can correctly keep the ethical principles as a decision-making agent with its embedded intelligence, and how the relative effects of alternative preferences and fields can be differently assessed when making automatic driving decisions. Furthermore, the ethical implications are critically evaluated at the individual, societal, and global level relating to these real-world examples, using various machine learning algorithms and driver-based configurable and real-time modifiable parameters [5].

Due to advances in AI and Machine Learning(ML) sectors over the last decade, multi-faceted autonomous algorithms have been introduced to various digital ecosystems with promising results. With the implementation of ML algorithms, tedious human-driven tasks can be automated, data-driven decisions can be enhanced, and smarter products can be developed. Considering these advantages, integrating AI and ML techniques into autonomous digital systems has become a trending application space. However, it's worth mentioning the significant ethical implications of using these technologies when operators and policy-makers apply them to critical decision-making systems. Consequently, a widely popular topic relating to ethics considers how these ML-powered AI operators make critical decisions, and what roles they will play in digital societies while making these crucial decisions. Specifically, in recent years, the gradually increasing interests turn toward smart interveners in the digital decision-making processes like self-driving cars. As a suitable scenario, we'll put forth how an autonomous vehicle intervenor (AVI) employs its ML algorithms to make critical driving decisions and the subsequent ethical concerns that rise due to these interventions [6].

2. Foundations of Machine Learning in Autonomous Vehicles

In fully self-driving vehicles, the human machine interface will become mainly speech-based given the drivers will no longer be required to monitor vehicular controls or the environment. The driver's intention and destination must be recognized and interpreted by the vehicle and the expected vehicle behaviour within the traffic context must be inferred and communicated to other traffic participants like pedestrians and other vulnerable road users. Any negotiation must be done transparently, explaining and justifying actions undertaken. These requirements demand more natural, intuitive and efficient human machine interface systems. [1] Services for conversational agents, sight tracking, and auditory assist systems is a strikingly advancing field and nowadays becomes developed in various technologies. These modern, but still under research, systems can help autonomous vehicles to answer the abovementioned challenges and can help to predict driver's destination in order to adjust the onboard in-vehicle infotainment system settings during the journey. Similarly, vehicle to pedestrian and vehicle to vehicle interaction related services are also needed to improve all involved road users' safety, allowing them to understand voiced communication and displaying context-related messages. This current study presents CREMA (Conversational REasoning in Memory-Augmented architectures), the framework for integrating context-aware multimodal conversational agents into autonomous vehicles.

[7]Machine learning has a particular significance in autonomous vehicles even though it is a subset of artificial intelligence. Moreover, the application of machine learning develops in an accurate response to surroundings and real-time quality for autonomous vehicles by using improved code learning. The demand for analyzing ethical autonomy, self-determination, and data privacy for developing collaborative autonomous driving decision-making modules are the need for National Research Council (NRC). There are many impacts or implications evolving in transportation-related societies by using autonomous vehicles.

2.1. Overview of Autonomous Vehicles

Reinforcement learning is a learning technique by which the agent learns through the measure of rewards and through environmental feedback. Reinforcement learning based on the maximum reward rate, the agent makes the best decision for the current state through taking actions. Reinforcement learning has been applied for two types of this type of learning, which are Exploitation, and exploration. When an agent behaves according to the policy, it is known as exploitation which has currently learned and optimization of the policy by which the agent is known as exploration. In the future when, AVs are moving goods so that the control of the AV is even more important for the economies of the world, AVs might need to move without a driver, and they can even learn and improve operation by combining the global operation and learning, through AVs connected to other vehicles. The automation of A/B/D Pillars offers sight driving zones, which will probably jeopardize the interactions between AV and human driver, such as recognition of the driver intent and implementation of Human Machine Interface.

[8] [9] A level 0 autonomous vehicle (AV) is defined as an Assistant system, meaning that it requires human intervention to achieve 100% autonomous operation, while level 5 AVs, on the other hand, do not require driver intervention and can operate in any environment, and

on any road type. They must all rely on deep learning algorithms to obtain the data, and to make decisions or predictions. AV learning algorithms are classified into three primary types: Supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is based on a technique that trains the model to learn from training data, most of the learned algorithms fall under supervised learning. The training dataset has several resources, each of which includes paired inputs targets. AVs have to navigate based on environmental inputs, such as traffic light detection, crosswalk detection, lane detection, and obstacle detection, in unsupervised learning.

2.2. Fundamentals of Machine Learning

The possibility for a vehicle to be fully autonomous raises the question of ethics of these systems and automakers in the event of an accident. Studies drew a need for the vehicle decision-maker to be transparent about its decision logic and ethically sound. The possibility of comparing the behaviours of different vehicles is also seen as a way of promoting respect for the principles of transparency and respect for individual preferences and consequently as promoting ethical behaviour on the road. The knowledge of these algorithms can eventually lead to effective design of the moving environment so that the data that the decision-making systems receive is more appropriate for their decision-making. Machine learning techniques, when explained and implemented, show promise for overcoming formally induced biases and building a design that can take better account of the user preferences. Thus, the implementation of machine learning could be ethically and legally accepted on the road [10].

Machine learning (ML) used in the decision-making process of autonomous vehicles is of concern from an ethical standpoint due to the potential consequences of these decisions. As such, autonomous vehicle technology should be approached from an ethical standpoint ensuring that a lesson is learned from the development of vehicle technology [11]. In 2011, Google introduced the fully autonomous car, and since then Automated Driving Systems (ADS) have been developed around the world. In the event of a major accident using vehicles with ADS for their decision-making software, the responsible driver as well as the involved companies would require, among other things, to document, explain, and justify how the car's decision-making software weighed the ethical considerations of the current situation. Therefore, addressing ethical issues in the development of autonomous vehicles is a primary priority [12].

3. Ethical Frameworks in Technology and AI

Moreover, this study evaluates, while virtually demonstrating, the ethical issues on sample personalities branch and determines all equivalent Klipper, collectively together. Also, the collection of attributes is classified into four subcategories called 'Corporate', 'Ethical', 'Explanative' and 'Utilitarian' and the relationship between vehicles carrying box and cube in this classification is determined from four different ethical systems at theoretical and practical depth. As a result of this categorization process, it has been revealed that consumers who make the statement were mostly applicable in the categorical ethics component. Virtual machine learning algorithms with automatic notes are suggested for automobile. It is essential to provide information on some topic like human behavior, physical condition of the driver and the state of health [13].

Amongst other distributed ledger technologies, permissioned so-called Federated Learning (FL) has introduced by the Google through a permissioned (an emphasis unlike the Google) FL application, which is permissioned and its security levels are regarded as crucial for the involved stakeholders, for sharing models without sharing raw data [7]. Because the billing system is more suitable for FL than Permission Blockchain. Smart contracts provided by the Ethereum permissioned blockchain have found to be less convenient for handles with the legal requirement. The focus of this paper is examining the current legal framework within the road traffic and road traffic insurance system. The entirely decentralized system directly verifies and automatically optimizes the self-driving cars with Fiat. The suggested procedure also examines the parameters in 20 input glare dataset approximately 50 times the normal visit, making a similarity 35% of dataset sample [14].

3.1. Ethics in Technology

In AI systems—like automotive systems—the restriction to algorithms is particularly challenging because AI systems always already contain the societal and normative conditions under which they were developed and used. There is no "normal" or ethical AI, AI can only be developed to mirror our ethical norms and moral values (related to these norms). To grasp what ethical norms govern AI development and use, we discuss technical and engineering literature as well as political and ethical documents. Uncritically, autonomisation—and therefore the digital qualification of society— poses challenges to the self-conception of democracy. [15] The requirement for a "lawful AI" proposed by governing bodies relies upon these challenges, as does the suggestion for future, ethical AI developments, put forward by the European Commission.

Trustworthiness is a cornerstone in digitalization and specifically in artificial intelligence (AI). Trust must be ensured by addressing questions of purpose, transparency, responsibility, and security. For ethical AI design, the ethicality and societal effects of AI must be analysed, as well as the subsequent positioning of the AI system in a social context such as work. With the improvement AI and machine learning algorithms, systems and machine learning methods become increasingly complex and opaque, strengthening the role of careful ethical considerations during the whole system lifecycle. This includes transparency to prevent a "black box", making decisions more understandable and traceable. This includes getting a better understanding of AI's influence on human dignity as well as addressing distributive justice along the whole value chain of AI. The crucial issue of control is addressed by asking for a citizens perspective for e.g. smart-cities planning. [16] [17]

3.2. Ethics in Artificial Intelligence

Despite being a set of technology and codes, AI still encompasses cultural influences. Every decision-making from AI technology impacts society, culture and its developers. The discrimination in data types is inherited within the machine learning models too, which implies the implicit bias present in data will be also there in the prediction model. Some of the built-in biases might be too subtle and difficult to find, but all can destroy the model's performance. In conclusion, to combat the biases stemming from this kind of data collecting approach in case of machine learning, a bi-directional approach including 'Explainable AI' and transparency and including general R&I perspective for autonomous innovation projects are suggested as good practice to facilitate the entire development cycle of autonomous technology (Burton et al. 2018: Sanyal, 2018). The view from the RRI perspective focuses on holistic approach along entire innovation process to ensure the uptake of the technology results in trustful and acceptable innovations by providing agile pathway to integrate mature, secure and privacy protected innovation (Bengtsson et al. 2019). This will serve to ensure that innovations meet social needs and are in line with social values [18].

Machine learning is at the very heart of artificial intelligence. It has to gather its own experiences and learn how to use them, hence it can also make mistakes. In addition, machine learning can also generate biased solutions. Depending on the available data, the learning phase could generate inaccurate solution. What is peculiar about machine learning among the AI techniques is that not only biases of developers, but small differences in the data can have a crucial impact on the final solution (Kamishima et al, 2012). The lack of transparency and interpretability (of the learning procedure) are a problem when we consider the impact of these AI techniques on end-users. To ensure a degree of fairness towards individuals and to avoid such misadventures, we propose to rely on more deterministic and non-ambiguous approaches, such as formal ethics [17].

4. Machine Learning Algorithms in Autonomous Vehicles

In the case that AVs need to perform an emergency action, the communication of its intention to the surrounding traffic is highly relevant. If AVs constantly behave conservatively and brake at any uncertainty, they most likely will get involved in rear-end crashes and cause traffic congestion. With such a pessimistic policy, it will be hard to gain the trust from the other traffic participants in actual mixed traffic to ensure a smooth takeover procedure. Consequently, AVs should only perform emergency actions, like a sudden change of lane, if no longitudinal control can avoid an accident and if the resulting trajectory is clearly communicated to the ego vehicle's surroundings. Nonetheless, during high assertiveness, that is, driving with high precision, it is especially challenging for AVs to quickly and safely convey its intention."

"Machine learning (ML) subsumes different techniques that, instead of being programmed to follow rigid rules, are trained on large databases of examples to analyze input and provide an output [19]. In autonomous vehicle (AV) decision-making, ML has certain advantages over classical computer vision algorithms, and primarily concerns two aspects. A first branch of works deals with the improvement of the detection during emergency braking of other vulnerable road users, given that it constitutes a particularly crucial traffic-situation for the takeover of AV control by the human driver [20]. A second line of research concerns the application of machine learning to improve the communication of AV intentions to other traffic participants [7]. In line with the collaboration with a team from Audi, this paper particularly focuses on the second topic.

4.1. Types of Machine Learning Algorithms

In supervised learning, a learning algorithm learns to map input data to a label associated with it [8]. Based on how the training data for learning is given to the model, supervised learning algorithms can be categorized into two approaches: regression and classification. In regression, the output variable is usually a real-valued scalar. In autonomous vehicles, predicting velocities, steering angles, etc. can be modeled as regression problems. In contrast, in classification, the output is a discrete set. Detecting objects in a camera and predicting the presence of an object that requires a binary response (e.g., presence of pedestrians, presence of a stop sign) can be modeled as a binary classification problem. Some other objects such as bicycles, motorcycles or cars can be classified under multi-class clossifications.

Machine learning (ML) algorithms can be classified into three categories: supervised learning, unsupervised learning, and reinforcement learning [21]. In supervised learning, models learn how to predict the output for a given input based on data where the output is already known. In unsupervised learning, there is no known output associated with the trained data, and the model learns to make predictions about the outputs. Reinforcement learning is different from the other two approaches, in that, the model receives feedback based on the actions it takes and learns to predict action sequences that can ensure success in a specific task. Among these, most of the recent works focus on supervised learning due to ethical implications related to interpretability and transparency in predicting autonomous vehicle decisions and outcomes.

4.2. Challenges and Limitations

A basic unsupervised model can be exploited by an adversary to inject data to cause identity spoofing. An insider adversarial attack with physical, digital, or human authentication rights can be fatal. It hence calls for optimal design, development, deployment, and control in the digital transformation of the modern world to assure ethical autonomous driving. Such challenges intend to establish and comply with systems ethics, the principles underlying individual and societal morals in engineering decisions related with automated systems and their impacts on society. Hence, security vulnerabilities and privacy related concerns remain the principal challenges which arise due to the usage of deep learning and analytical techniques. But, the solutions to these problems can be created and adopted so that the proponents of the technology can enjoy its privileges without creating issues for society [2].

[20] Despite the benefits that machine learning algorithms can provide for autonomous vehicles in terms of their decision-making, the adoption of these algorithms can carry ethical

implications that must be considered. However, machine learning algorithms can severely impact ethical constructs due to challenges such as discriminative and biased decision-making [22]. Many other technical challenges surround these decisions as reported by Wang and co-authors. Moreover, Deep learning models sometimes suffer from model hallucination, update poisoning attacks, vulnerability to adversarial attacks, and lack of warranty proving. However, by simply employing deep learning and analytical techniques to collect and analyze data, we can not only avoid ethical issues in autonomous driving, but also promote it.

5. Ethical Considerations in Autonomous Vehicle Decision-Making

The utilization of AI for designing and controlling semi- and full-autonomous vehicles is current subject of interest in various research fields. Also, autonomous AI-based systems but human operator and monitor system are discussed in the related current study. In one of these research papers, we discuss ethical issues and privacy concerns for several machine learning algorithm-based applications in computer vision, robotics, self-driving vehicle, and reinforcement learning-based smart grid, etc. Ethical dilemmas in emerging technologies, especially connected to the autonomous vehicle, become a controversial topic. In the first part of this chapter, the difficulties and analysis of whether autonomous vehicles can make the right decision by facing ethical dilemmas (such as, between selecting killing of the single pedestrian and saving five crashes, or bringing to starboard to crash tree or to bay to stop a crash (faulty drivers should crash into hazards of open set rather than the surrounding environment to save more lives)) of autonomous vehicles are considered and solved, as well as the public opinions on which situations and preferences will be highlighted.

[5] [19] With the advancement in artificial intelligence and robotics, significant power in decision-making is converted from humans to machines. AI presents machine ethics, which investigates how to design morality-compliant algorithms and systems. When considering implementation of ethics in autonomous systems, characteristics of AI should be considered, such as reliability, autonomy, and accountability, pillars of robust ethical AI. Most of the ethical dilemmas relate to autonomous vehicles' decision-making during extreme and rarely occurring traffic scenarios. This topic is especially focused on the application of autonomous systems and AI in automotive driving in order to highlight the severity of the ethical implications of AI in real-life scenarios. These cases also relate to the limitation of AI and

decision-making limitation, which brings challenges and impact on human-level fairness, explainability, transparency, and public policy application.

5.1. Safety and Risk Assessment

* Allowing risks when the AV could not have a NASH equilibrium solution, * In case of uncertainty, maximizing the safety of several types of external facilities (pedestrians, cyclists and cars). According to the authors, in future AVs, they will need to perform much more socially acceptable decision–making.

To avoid incidents or accidents, AVs will have to compute which value has to be minimized given the current constraints. The authors then emphasized that algorithms could prioritize the occupancy of the AV, or on the opposite side the external facility, such as cyclists or pedestrians randomly visiting the street. Consequently, the authors proposed a new type of AI, named Moral Driving Torcs, capable to mimic human moral while computing orders in our Icarus simulation framework. They also showed that normal human behavior can be represented via:

Algorithms in autonomous vehicles (AVs) have the ability to make decisions while relying on risk and safety evaluation [22]. In the MAKIS project, an algorithm was introduced as the main module in which intelligent agents can communicate and verify each other to avoid collisions. The nonconformity between the actual environment state and the environment as calculated by the algorithm increases when the sensor system in the Autonomous Vehicle (AV) fails. This makes the legibility of the possible obstacles inaccurate, and the higher the inconsistency, the more unpredictable will be the AV behaviors and actions in the system, which spontaneously results in a safety risk. The authors argued that the use of deep learning and machine learning in autonomous vehicles has significant problems of equity, fairness, and the ability to understand the appraisals made by algorithms. In current AVs, the output of the machine learning model determines the next step completely.

5.2. Justice and Fairness

In a paper published in the Journal of Experimental & Theoretical Artificial Intelligence back in 2011, the authors of the Trolley Problem — a thought experiment that famously asks whose life an autonomous vehicle should prioritize in an emergency — suggested that moral judgement depends (in part at least) on inferences about drivers' "interactive competencies".

Training your car to distinguish vulnerable overprivileged bystanders is, in other words, moral progress; explicitly coding this distinction means taking a shortcut whose logical conclusion is an AV that refuses to halt for people crossing the road in a wheelchair. "Governments should ban default settings that prioritize identifiable groups," they write. A life saved is a life saved, regardless of whether you know anything about the person who will be doing the saving.

Ethical considerations are an essential consideration when designing machine learning algorithms for autonomous vehicles in order to avoid building in systems that inadvertently promote harm and bias. One term that has seen an uptick in usage in the transportation context is "external governance" geared toward ensuring that the development and use of technology is aligned with social values [11]. This problem also has implications for fairness: if autonomous vehicles disproportionately injure certain groups of people, in the absence of alternate means of achieving justice this could be seen as equivalent to an unjust inclusion in an unfair social contract [14]. The same problematic logic can be applied to privilege, if life-saving vehicles are designed to favour certain groups of people, they may be ultimately scapegoated for the social structures in which the preferencing algorithms were nurtured. "In the absence of other legitimate concerns, privileging one group can be seen as unjust".

5.3. Transparency and Accountability

In general, thanks to these methods it is possible for a developer to retrace the entire decision process, understand a malfunction which determined an illegitimate choice and intervene to change the parameters of the model or produce better data. At the end avoid the occurrence of analogous events in the future. Moreover, it is crucial to guarantee not only the transparency and openness of an intelligent system but also the system accountability. Accountability requires that a system of responsibility is created and it is determined which entity is going to be responsible for actions realized by AI. This has specific relevance in the case of accidents and in general of bad actions.

Transparency and Accountability: In the outcome of any negative impact of decisions by a robot, autonomous vehicle or an AI based decision-making system, it is fundamental to understand the logic that the system had used in order to generate it [5]. If someone gets hurt because the car swerved, it is vital to understand the necessary steps that the objective function has moved through and it is essential to understand where it has gotten lost in order to return to a positive ethical logic [23]. Hence, ethical principles require that the logic of such systems is accessible and understandable. How can transparency be given in a concrete way? Each decision coming from an AI is the result of a function called model [8]. This function has seen as an input the data information and has produced an output. Hence, to understand a decision, it is necessary to understand the parameters of this model and the weights of this model, because they have to describe the model. Additionally, transparency could be given by opening and making accessible data sets used to train the decision systems, so that the origin of the elements that have contributed in taking a decision could be retraced.

6. Case Studies and Examples

Given the remarkable amount of theoretical papers that deal with ethical implications related to autonomous vehicles, we set out to investigate the current state of practising developers and automotive companies by conducting an online, quantitative survey with 120 participants from the automotive industry. Of those, 13% are exclusively automotive suppliers, 16% concern themselves only with manufacturing end-products, and 37% declare to be both. To our knowledge, the study at hand is the first to investigate what manufacturers and suppliers are currently doing with respect to the ethical aspects of machine learning algorithms for autonomous vehicles and their decision-making.

[7]There are a variety of ethical dilemmas could arise in AV decision making. In 2018, Gesualdo, Schumann, and van de Poel [vande Poel, de Winter, & Happee, 2022] listed, among others, moral and legal responsibility, privacy, and moral decision making in critical incidents as the focal ethical issues. Fleischer [Fleischer, Salge, Hussain, & Sendhoff, 2018] drew upon Fedorov, Kozarev, Rozhkova, Kiseleva, and Sazaruk [Fedorov et al., 2017] and Szollosi & Hernandez-Lallement [Szollosi & Hernandez-Lallement, 2017] to outline and analyse a number of ethical considerations, including distributive justice, harm, bias, social value, and non-maleficence. Research on the ethical implications of driving decisions concerned with moral dilemmas includes, for instance, autonomous vehicles weighing either self-value and utilitarianism or flexibility trade-offs applied to ethical principles [Gogoll & Müller, 2017], and the potential impact of changes in traffic law on the interpretation of rights and obligations for automated vehicles [Jeppsson, 2019]. Research has also touched upon broader and conceptual theoretical aspects of AV ethics and morality, including ethics and responsible AI deployment [Andersen et al., 2018], [9] moral dilemmas for moral machines, which are used

to assess the ethical performance of new algorithms, and why mainly ad-hoc legislative responses will fail in the overspeeding race of bringing AI/automated driving products to market [Liebsch, 2017], [19].

6.1. Real-world Scenarios

In this section, we will present six real-world scenarios to illustrate the difficulties an autonomous vehicle will face when it needs to make a choice on who to harm in the case of an unavoidable accident. These scenarios are neither hypothetical nor rare in actual traffic, and it is crucial that automated vehicles can fulfill their duty in these kinds of everyday traffic situations. Rather than being less likely to take place than other scenes, some of these scenarios could actually have occurred at the time of writing. Hence, log data of German police records were used to collect realistic scenarios where an autonomous vehicle (L3, L4 or L5) had to make a decision on whom to harm. Even though the current work does not aim to propose solutions, it seeks instead to increase the public's awareness about the challenging scenarios, especially prior to the large-scale deployment of automated vehicles. [2]

The behavior of automated vehicles in critical situations poses ethical questions for both regulators and designers. When an accident in inevitable, which action should the vehicle choose to minimize harm? Moral philosophers face the same question when they imagine you and a group of people standing on a railway track with a trolley heading their way. By shifting the trolley's course, you can save your group by sacrificing one person. Foot (1967) first introduced this dilemma. Its moral relevance subsequently inspired a vast amount of literature on the topic, sometimes referred to as "moral puzzles" [24].

6.2. Ethical Dilemmas

Though AVs hold an advantage over humans in analyzing vast number of potential outcomes and making split-second decisions in unexpected traffic situations, it raises an ethical debate—wherein the expected moral norms for AVs are found to differ from person to person. As a consequence, one of the greatest challenges in the deployment of AVs is an ethical one: how to fairly distribute the risk of accidents between various parties such as passengers, pedestrians, and/or cyclists [25]. In order to address this concern, researchers have questioned the deployment of AVs through models developed by querying people with various moral dilemmas known as "moral machine".

Recent advancements in machine learning algorithms for autonomous decision-making capabilities have bestowed upon these vehicles human-like abilities [26]. Decision-making under conditions of uncertainty necessitates the anticipation of a limited number of possible outcomes for each possible action. Utilizing advanced artificial intelligence technology, autonomous vehicle systems can aver the outcome of each action through scenario planning based on prior learning trajectories which are often developed to optimize minimization of passengers' mortality risk [27]. Thus, the autonomous vehicle (AV) will display a supremacy even in all adversarial traffic scenario through time-dependent high-level decision architectures based on deep reinforcement learning.

7. Regulatory and Policy Implications

The need to consider the ethical implications of AV development and deployment in concrete regulatory terms has also been a theme in recent debates that emphasize the importance of normatively sound regulation over and above universal guidelines. In the context of ML and AI ethics in medical design, prima facie ethical principles that may need to be complemented with situational responsibilities and responsibilities toward existing institutional structures have also been emphasized [6]. Fostering human rights and addressing risks to fundamental rights and values, in turn, are also a core part of the European Commission's efforts to develop a comprehensive AI policy framework. Therefore, concrete expectations about equivalent safety in AVs and fundamental rights emerging from other legal frameworks, including those that protect privacy, are also important elements of any ethics regulation [11].

The increasing interest in both the safety and ethical aspects of autonomous vehicles (AVs) has led to the development of a number of guidelines and campaign. Given the potential impact of AVs on road safety, discussions concerning the ethical implications have been particularly relevant. Nonetheless, there is also no shortage of views that suggest comprehensive regulatory reform in this area. Furthermore, bioethical AI guidelines like Ethically Aligned Design [4] similarly emphasize ethical principles embedded in AI systems as background norms, while suggesting that a shift in thinking about AV decision-making is critical. Such guidelines also imply that the regulation of AVs should be primarily principles-based, with a clear consideration for human values and ethical commitments.

7.1. Current Regulations

To facilitate a sequential implementation of ADMS, a strategic plan was established by the Department of Automotive Engineering, at the Chinese-German Institute for Legalization, Faculty of Automotive Engineering, Changchun University of Technology. ADAS can reduce marginal fatality risk (the fatality risk while being under full control without ADAS) and severity of crashes (in the case a crash occurs) and could start implementing in a time frame of five years. To comply with the social standards, ADMS are required to be able to interact adequately with all road users (including pedestrians and cyclists) and social goods (including ecological, energetic and social effects) driven by ethical decision making models. Achieved by an integration of DMS, ADAS and autonomous driving functions, ADMS can increase both the level of traffic safety and mobility [28].

input_text: The public discourse often has oversimplified views on autonomous vehicles (AVs), and the complexities of ethical and moral decision making in AV driving scenarios is not well understood by the general public. Considering the legal and regulatory implications of ethical decision making for AVs, the question arises whether social, ethical, and legal standards are up to date and sufficient, and whether additional guidelines and regulations would be useful or even necessary depending on the moral scope. Currently, existing legal regulations from all UN member states reflect a human centric perspective with heavily human centered regulations, which may be under-specified or even absent when it comes to AVs, may limit the possible decision making spectra of AV algorithms, and may not be sufficient to govern the complex ethical situations AR coud encounter [9].

7.2. Proposed Policies

To this end, a 'policy-priority based' approach can be proposed for pedestrians (humans are considered before machines) and a 'maximum-utility based' approach can be proposed for animals and other vehicles (machines are considered before animals and machines) [22]. In this regard, the aim of this 'policy-priority based' approach is not about which member of the group to save but to ensure the rights of pedestrians or other roads users using priority-units for pedestrians and other vehicles. In traffic situations involving more than two casualties, the current ethical principles (utilitarian/non-utilitarian) are replaced by the ethical principle (minimizing damage) on which different stakeholders find the common ground to develop the solution. Thus, it is proposed that there be a move to a priority-based approach to solve the classical trolley problem rather than a maximum utility-based approach.

One option to mitigate the moral dilemmas that arise in traffic accident scenarios is a 'principle of necessity' [14]. This principle postulates that in critical situations, that is, if the accident cannot be avoided, the automation systems must act to minimize damage in a way that is ethically justifiable. Another way to address the problem is to implement programmed rules that are based on legislation and morally acceptable ethical considerations [7]. The problem with these two approaches is that there can be an infinite amount of unforeseen scenarios/harmful outcomes for which the correct ethical choice may be impossible. Accordingly, a more efficient approach could be to implement a responsibility-sensitive-safety (RSS) framework, which is based on ethically acceptable stated rules for traffic situations. The difficulty about the RSS framework may be the precise formulation and dynamic changes in traffic legislation.

8. Future Directions and Recommendations

To address this question and give practical recommendations for the ethical use of machine learning models in autonomous vehicles a mixed DL algorithm using interpretable AI should be developed. Further theory-based factors can be explained by a psychological factor: deontological business ethics prioritize moral duty, morality and social norms, respectfulness and fairness, while exclusivity (which was actually very small) overrule safety concerns, helping with the conclusion that the best algorithm decisions should respect social norms and dignity [9]. Systematic empirical testing in this vein would be instructive. Furthermore, the role of psychological distance should be further explored; we suggest that our in-group bias responses have possible practical implications and note that previous empirical queries have suggested that the age of targets in hypothetical moral dilemmas may play a key adversarial role.

The proliferation of machine learning algorithms and conversational agents makes it imperative that attention be paid to the design of algorithms that conform to ethical norms. The results of our study show that the ethical implications of machine learning for autonomous vehicle decision- making are severe, and the fairness and morality of AV algorithmic decision-making effects have important implications for the ethical design and deployment of AV algorithms [3]. Indeed, we believe that ethical machine learning for AV algorithm development should be a key research priority. Approach: Future theoretical and empirical work (both qualitative and quantitative) could explore the favour factors, their use by algorithm creators, and the ethical, philosophical, and policy implications of the objective-based reasoning. This is a key issue because the Chinese Code firmly state that when in conflict, the AV should prioritize priority of the lives of passengers above all other considerations [7]. Our hypothesis tests allowing for this dynamic are as follows: H1A: A passenger impact involving a carload of people will be seen as more ethically appropriate when the car has a preference for either young or elderly passengers; H1B: A pedestrian impact involving other factors besides the number of lives being saved will be seen as more ethically appropriate when the car has a preference for either young or elderly pedestrians.

8.1. Emerging Technologies

Neuroscientific approaches (e.g. Grau et al., 2003) suggest that rodents could serve as a suitably complex test animal as their mental and sensory abilities are considered to be relatable to the predictive processing notion. As of today, they present a paradigmatic case study of how to implement the ethical concerns of the machine learning of AVs in a particular environment full of different kinds of intelligent behaviors, being it decision-making or just plain impulsive trouble-making. Hence, in the field of predictions for normative behavior prompted by robots, autonomous cars have become lovable pets of ethical reasoning [22]. Similarly, in Europe, the debate of ethical questions in machine ethics and robotics has started to address AVs. Galperin (2013) presents a directive-conjunction approach which aims to realize a balance between principles of moral ethics in such a way that these can be put into implementation too by means of machine ethics.

Algorithmic decision-making in autonomous vehicles (AVs) will have several ethical implications as AI continues to embed in the emerging AV technologies. Besides the role of the machine learning algorithms in AI in determining who to sacrifice in unavoidable accidents, and whose lives to prioritize, it should also be asked to what degree such "ethical reasoning" might infringe on traditional legal and moral responsibilities for AI-planned actions. While in the case of social media bots with concrete personalities, encouraging particular interactions with others and communicating in specific ways is the desired objective, an ethical autonomy designed to regulate the (non)autonomous car's behavior will have to satisfy a much more complex set of norms, pedestrian traffic swarming being one of them [29]. Specifically, the real challenge is to overcome the disruption to normative traffic regulations by targeting socially centered legal monitoring on the possible system permissions. VW (2016), for example, describes this scenario indirectly by focusing on the application of ethical norms and culturally reflected ethical values on socially intelligent autonomous robots/personal data robots.

8.2. Ethical Guidelines

Even if it were reliable to achieve reliable automation behaviour, manufacturers would of course still have to align their policies with three additional general ethical principles of the European Commission: to explain how an autonomous vehicle comes to a decision; not to take technical action behind the back of the human driver; and make sure that unintentional decisions by autonomous vehicles are as 'friendly' as possible. It is noteworthy that Gordon et al. are assuming that moral liability (positive—the decision-maker is allowed to assume more severe consequences; negative—the developer/manufacturer is allowed to assume the resulting consequences) always lies with the human and not with autonomous agents, probably based on the observation that we have so far seen only acceptance of highly observable failures in human operators governing technical systems. This situation will change dramatically, however, as soon, as we gain trust in machines being superior to us in terms of safety.

The ethical implications of machine learning algorithms for autonomous vehicle decision-making have received increasing attention in recent years. Ethical challenges stem from uncertainty about how autonomous vehicles will behave in extreme traffic situations and uncertainty about the ethical choices that will be pre-programmed by OEMs. Responsibility for particular ethical choices should therefore be shifted from machine learning onto decision-makers as far as possible [14]. "Ethical guidelines for a trustworthy artificial intelligence", which were published by the High-Level Expert Group on Artificial Intelligence of the European Commission, should be applied to machine learning algorithms for autonomous vehicles. More specific ethical principles and guidelines are also needed in order to resolve ethical issues that are unique to autonomous vehicles. For example, for machine learning autonomous vehicles, reinforcement-learning algorithms may make choices in certain situations that are not in the best interest of their occupants. It is clear that it is necessary to develop the best possible risk assessment and risk minimization strategies and to work on ethical standards so as to ensure the basic safety of decision-making in humans and sentient agents—the decision-making liability in this case is partly or entirely transferred from the developer to the decision-maker.

9. Conclusion

The ethics problems of autonomous vehicles have been studied for a long time. The main existing problems are how to deal with accident scenarios, that is, how to make a choice when facing various accident scenarios, and the ethical problems when experiencing system failure. Multiple care should be taken when autonomous vehicles (AVs) are tested and/or put into practice in an urban environment, due to the shared public funded spaces and our need for safe and trustworthy systems. 1. as the sharing of the traffic space between pedestrians and the AV; 2. how to personalize on-the-spot and infrequent decisions; three possible reconciliations with the trolley problem. The XR-connecting of AV will have emotional, political and societal effects.']].

[[article_main_idea] This paper has examined ethical implications for future autonomous mobility specifically subway and tram traffic, schools, pedestrians and extended public, partnerships also when assisted by machine learning algorithms and tight and loose ai. The conclusion is massification of the technological driving system can spans generations and includes ambiguity for how to inform passengers about the trustworthy of ai. The development of autonomous vehicles (AVs) makes it possible for people to go anywhere without encountering human-driven accidents at the definition of an autonomous vehicle (a). A self-driving car is a car that can travel to a pre-set destination without a human driver. Person a can pull out the phone, chat with someone, and even work in the car. The concept of autonomous vehicles emerged as early as the Coca-Cola Automatic pilot, piloted by a sensor technology sponsored by Coca-Cola in 1953.

To forestall these and other (mainly in the human lens non-understandable) mischiefs, three solution approaches are brought forward. One is based on a retribution-entailed framework of compatibility, the policy of respect. The second makes use of soft engineering in an introducing or enabling manner, with the notions of intelligent machines (IM) and setting-in-stone (SiS) steps. The third approach, as an overlapping and fill-up solution, respects the ethical arc of an action theory framework and proceeds according to Himes' work. Last, it is sketched that quart two of solutionary ethical attention can be easily and comfortably completed (or at least helped) by tools, developed and filed under the headline artificial ignorance. The tools will feed human user-friendly picturing to the key stakeholders, who are lay(readers) in the primary application context(s). No matter which of the differentially coloured ethical preparedness approaches one opts in a development boom of autonomous vehicle systems, no PKI of them can ignore the fact of oppressed lay human acceptability stakeholder decisions at the end of the day. With so many diverse stakeholders, the decision-making process in the Council of the European Union appears to be a very encouraging 'people encapsule'.urança de sistemas automotivos. " Tese de (doutorado). Instituto Politécnico da UFRJ, Rio de Janeiro, 2021 article_id: 68afc66b-46c1-495b-9b80-a6b4a0e93049 article_title: Ethical Implications of Machine Learning Algorithms for Autonomous Vehicle Decision-Making article_main_idea: This paper has explored ethical concerns regarding autonomous vehicles from the perspective of machine learning (ML) algorithms. The focus was on discussing implications for the future and associated social and policy ideas.

Implementation of autonomous driving in current road traffic requires safe, efficient, compliant, and user-acceptable development processes and technical prerequisites undergoing verification. It is not safe (also) because it creates and requires 'social licence'. In extreme cases, this includes permission to foresee—implicitly or explicitly—sacrificing pedestrian lives to save oneself or one's loved ones or traffic dilemmas ending in the well-known so-called trolley problem. While this problem in philosophy seems as old as Thomas Hobbes (1651), the publication of it at large for and in the automotive world is considered to be due to Joshua Green (2008). Here it was loss of lives at the hand of automated vehicles—hopefully 'only' as extreme reasoning experiment—resulting in highly undesirable algorithmic consequences without ethical supervision.

[[article_main_idea] Every year, the number of autonomous vehicle (AV) trials and deployments in urban environments has been increasing. With increasing cooperation and urbanization trends, there are continuous reasons for the AV to operate in the urban environment. This paper highlights the major ethical controversial decisions an AV should make in urban mobility. Within our framework, the AV is mapping low-level sensors to high-level information about the urban environment. Multiple complex sensors of the AV in a setting of high-uncertainty, -complexity and -politicalness result in a need for complex action-making systems; one of the systems is a paradigm shifting shared metric for AV control and reinforcement learning. Sub-goals are argument strategies for AVs, or waxing the trolley problem.

9.1. Summary of Findings

The driving scenario used for this study can be described as a T-intersection where the AV is coming from the bottom and has the road top as its main road. Straight on, it has the possibility to avoid a collision with a motorcycle. The width of the top road causes a the motorcycle on the top road to be hidden from the worldview of the AV until the conflict point. In the straight, the visibility is obstructed by road buildings up to the conflict point causing the AV to have to take its decision based on prior knowledge. The activations of AD networks can be transferred to the law discussion because the strict and involved handling of the concept of intentionality included in the crime definition. The network decisional parameters are combined with societally informed advanced causal graphs of law and reflect a classical shift in emphasis from rule to case-based reasoning.

Autonomous Vehicles (AVs) will somewhat be the future of personal transportation in the medium to long term [27]. Autonomous driving (AD) together with an in-depth description of boundary conditions, i.e., how the AVs are set up to behave in them, will entail that automatic decision-making/decision-support systems have a major impact on future vehicle related society debates [30]. Note that the number of decisions made by AVs will grow over time. The decisional power goes beyond utilitarian-peer related questions (that in worst case might lead to deaths or injuries of involved people) but also includes finding ways to secure resources and capabilities for societal activities.

9.2. Implications and Recommendations for Future Research

This chapter has discussed the complex set of rules that oversee AVs' decision-making processes, both drawing from the ethical considerations of implementing actual ethical AI and ML-embedded systems, showing that integrating machine ethics and normative methodological design into current discourses is an urgent need so as to avoid normative gaps as well as ensuring that decision-making processes are not a black box and are based on a holistic consensus. We envision a future when AI will be used not to automatize, but to assist and improve the safety of human driving. Moreover, we review some measures and potential policies that could help to foster an ethical advancement of ML in AV decisionmaking. Some other modalities are proposed, such as balance and stakeholder awareness, industry actor collaboration, or the prioritization of practices, criteria and methods. This allotment is shown by researching the shortcomings of machine-paced and machine-produced regulation and its subjective interpretation of valuing human life [1,4].

Current research has pointed out some ethical implications of AV decision-making processes that have been largely overlooked so far. One gap in the literature that our analysis points out is a lack of a clear ethical responsibility framework that calls for a multi-stakeholder collaboration with an explicit normative and evidence base. It has been argued that moral obligation needs to refer to the wider societal implications of AV decision making, which disrupts traditional human-agent interactions, raises new ethical and regulatory challenges, and transforms urban environments. Previous review studies have assessed lessons learned from these works and concluded that there are no clear and practical methodology development, showing the need for a deeper understanding of the meaning and implications of AV decision-making and AV ethics from a societal perspective or in how they shape human values in certain life situations.

Reference:

1. Pulimamidi, Rahul. "Leveraging IoT Devices for Improved Healthcare Accessibility in Remote Areas: An Exploration of Emerging Trends." *Internet of Things and Edge Computing Journal* 2.1 (2022): 20-30.
2. Tatineni, Sumanth, and Venkat Raviteja Boppana. "AI-Powered DevOps and MLOps Frameworks: Enhancing Collaboration, Automation, and Scalability in Machine Learning Pipelines." *Journal of Artificial Intelligence Research and Applications* 1.2 (2021): 58-88.
3. Ponnusamy, Sivakumar, and Dinesh Eswararaj. "Navigating the Modernization of Legacy Applications and Data: Effective Strategies and Best Practices." Asian Journal of Research in Computer Science 16.4 (2023): 239-256.
4. Shahane, Vishal. "Serverless Computing in Cloud Environments: Architectural Patterns, Performance Optimization Strategies, and Deployment Best Practices." *Journal of AI-Assisted Scientific Discovery* 2.1 (2022): 23-43.
5. Muthusubramanian, Muthukrishnan, and Jawaharbabu Jeyaraman. "Data Engineering Innovations: Exploring the Intersection with Cloud Computing, Machine Learning, and AI." *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)* 1.1 (2023): 76-84.

6. Tillu, Ravish, Bhargav Kumar Konidena, and Vathsala Periyasamy. "Navigating Regulatory Complexity: Leveraging AI/ML for Accurate Reporting." *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)* 2.2 (2023): 149-166.
7. Sharma, Kapil Kumar, Manish Tomar, and Anish Tadimarri. "Optimizing sales funnel efficiency: Deep learning techniques for lead scoring." *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online)* 2.2 (2023): 261-274.
8. Abouelyazid, Mahmoud. "Advanced Artificial Intelligence Techniques for Real-Time Predictive Maintenance in Industrial IoT Systems: A Comprehensive Analysis and Framework." *Journal of AI-Assisted Scientific Discovery* 3.1 (2023): 271-313.
9. Prabhod, Kummaragunta Joel. "Utilizing Foundation Models and Reinforcement Learning for Intelligent Robotics: Enhancing Autonomous Task Performance in Dynamic Environments." *Journal of Artificial Intelligence Research* 2.2 (2022): 1-20.
10. Tatineni, Sumanth, and Anirudh Mustyala. "AI-Powered Automation in DevOps for Intelligent Release Management: Techniques for Reducing Deployment Failures and Improving Software Quality." Advances in Deep Learning Techniques 1.1 (2021): 74-110.