

Ethics of Artificial Intelligence in Autonomous Vehicles

Sindre Bergsvik Øvstegård
Oslo Metropolitan University
Electrical Engineering
Oslo, Oslo
Email: s318337@oslomet.no

Kenneth Ramos Eikrehagen
Oslo Metropolitan University
Electrical Engineering
Oslo, Oslo
Email: s331475@oslomet.no

Abstract—

I. INTRODUCTION

The technology in automobile industry has made some remarkable progress the last century. They have managed to make safe, reliable and affordable vehicles. The last couple of decades have seen significant advancements in computation and communication technologies, and consequentially autonomous vehicles (AV) are now becoming a reality [1]. Several prototypes exist today. Among others, Volvo have started implementing autonomous trucks that deliver goods [2].

A. Autonomous vehicles

1) *What is it?:* A definition of autonomous vehicle (AV) can be a vehicle which senses its environment with little or no human input or interaction. This means that the vehicle is capable of gathering information about the environment around it such as objects, temperature, position, velocity etc. And execute tasks or functions according to certain directives.

2) *Why?:* Humans are biological lifeforms which depend and react on, among other things but not limited to, sleep, food, stress, isolation, social interactions and other similar processes. Besides being complex lifeforms humans react differently to these processes. Choices taken, in certain situations, differ greatly.

Machines today are not dependent on such stimuli and therefore they can, theoretically, continue with tasks to an infinite amount of time. At mundane and repeated tasks such as arithmetic calculations or retaining huge amounts of data, machines and computers excel with unrivalled precision and accuracy. This makes them good candidates for making pure rational and logical decisions based on optimal outcomes of certain criteria.

Based on this assessment machines would take more optimal choices over time in comparison with the average human. And their choices would be easier to determine based on simulations and testing scenarios. This would make outcomes and outputs deterministic and easy to predict and consequentially present a good basis for development and optimization.

II. THEORY

A. Technology

A brief presentation on certain technologies implemented in today's land-based vehicles. Both autonomous and non-autonomous.

1) *Real Time Operating System:* Cars today rely on numbers of sensors to acquire information about their surroundings. The velocity, GPS, proximity to obstacles, temperature etc. These sensors communicate to certain modules that are connected to an operating system (OS). This is often a specialized form of OS. A so-called Real Time Operating System (RTOS) makes up the central core and processor of the car. Within, all decisions based on data from the sensors are made and functions executed. Some aspects of RTOS are presented here.

a) *Hard time RTOS:* A deterministic RTOS. Meaning that the output of the RTOS should happen within a certain time or deadline. Used in mission-critical systems like medical devices, flight control space shuttles etcetera.

b) *Soft time RTOS:* Not as deterministic as Hard Time, but deadline of tasks and processes should be met within certain limits. The RTOS performance will degrade if tasks are never done within deadline and would be useless. Examples here can be audio-visual-systems for entertainment, games and general purpose OS's like Windows, Ubuntu and iOS.

2) *Camera:* Cameras are indispensable for providing visual information and feedback to any system, in regards to its environment. In comparison to radar, LIDAR or similar sensors, cameras are affected by weather, dust and pollution.

3) *RADAR:* Radio Detection and Ranging (RADAR) can complement the camera and is resistant to both pollution, weather and similar. It emits electromagnetic waves which reflect on obstacles and back to the RADAR. Distance, velocity and position (relative to the RADAR) can then be calculated with great accuracy. There are different types of RADAR and the most used is Frequency-modulated continuous Wave (FMCW).

a) *FMCW:* Allows the detection of small objects with the use of the Doppler effect. Since the known transmitted frequency is known, stable and slow-moving objects can be filtered out of the receiving signal. This reduces interference and increases the signal to noise ratio.

4) *LIDAR*: Goes under many names. One of them is "Laser Imaging and Ranging". This method utilizes laser beams and their reflection of objects to measure distance. Also called ranging. The lidar's output would be point cloud data, which provides all the necessary information, to a software for detecting and determining where potential obstacles and objects are located in the environment.

5) *GPS*: A property of the United States Government. GPS is acronym for "Global Positioning System". It utilizes satellites for positioning. The system is freely available for anyone with a GPS receiver and has global coverage. Several other systems from other countries exists. As long as a GPS receiver has line of sight to four or more satellites, a pinpoint of position can happen with an accuracy of two meters or less.

B. Autonomy

The definition for autonomy would be the capacity to make informed and uncoerced decisions, based on some set of information. From a more metaphoric perspective, this would be that an entity, person or organization would be given self-governing capacity. This to make decisions not forced on it by a higher ranking entity, regime, person etc.

C. Machine learning

Machine learning (ML) is a subset of artificial intelligence (AI) and by definition is "the study of computer algorithms that improves automatically through experience" [3]. Said in simpler terms it makes the machine/system learn from experience without being explicitly programmed, but to do so it need a lot of data. This data set is divided in to "training data" and "testing data", where the "training data" obviously is used to train the model and the "testing data" is used to check the accuracy of the model.

ML is often divided into three categories:

1) Supervised learning:

This type of algorithms is designed to learn by example, and the idea behind is that training this type of algorithms is like having a "supervisor" watching over the whole process. The training data consists of input that is paired with the correct output and during training the algorithm will search for patterns that correlates to the desired output. After training you get a model that takes new unseen input data (of the same kind) and determine which label this new inputs classifies as based on the training data. The whole essence of this kind of algorithms is to predict the correct label for newly presented inputs based on past data. On it's simplest form it can be written as

$$Y = f(x)$$

where Y is the desired output from the mapping function f with value x [4]. Supervised learning can be divided in to two categories

a) *Classification*: This algorithm is typically used to put the new data into a category or a class based on it's training data. Under training this algorithm will be given data points with an assigned category. The most common example is filtering out spam emails.

b) *Regression*: This algorithm is a predictive statistical process where the model tries to find relationship between dependent and independent variables in the data set. The goal of a regression algorithm is to predict a continuous number such as sales, income, and test scores.

2) Unsupervised learning:

In contrary to supervised learning that needs to train it's algorithms to recognize patterns in dataset with known labels, unsupervised learning is used on big dataset to recognize patterns in unlabeled dataset. In simpler terms this type of algorithm learns useful patterns and properties in the data structure in big datasets with no labels. The most used algorithm used in unsupervised learning is cluster analysis [5] [6].

c) *Cluster analysis*: Cluster analysis, also called clustering, is used to group datasets with shared properties in order to find algorithmic relationships. Each group made is called a cluster hence the name cluster analysis or clustering, but the name "cluster" can not be precisely defined. Since the notion cluster isn't well defined there are many different clustering algorithms out there, but they all have one thing in common they group data [7].

3) Reinforcement learning:

Reinforcement learning (RL) agent is learning by interacting with an environment, in other words it learns from the consequences of its actions. The RL agent can do this because of a numerical reward system that encodes the success of an action's outcome. The RL agent seeks to learn how to find the actions that accumulate the most reward over time. The RL agent learns from exploring it's environment and making decision upon how to act in every single scenario, instead of explicitly getting taught how to respond in advance. Thus this kind of algorithms is very complicated compared to the other two other ML categories algorithms [8].

III. IMPLEMENTATION OF AI IN AV'S

A. Which algorithms are used?

What kind of algorithm used is dependent on computational ability and determined on what kind of technology implemented within the AV. Be it GPS, LIDAR or similar. There is no problem implementing all sorts of algorithms, but a factor that limits range of viable candidates are, the beforementioned demands and also demands in regards to safety and security.

B. Safety

C. Implementation

D. GDPR

IV. CONCLUSION

REFERENCES

- [1] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, "Towards fully autonomous driving: Systems and algorithms," 2011.
- [2] P. Sawers, "Volvo's vera autonomous trucks will transport dfds goods on public roads." <https://venturebeat.com/2019/06/13/volvos-vera-autonomous-trucks-will-transport-dfds-goods-on-public-roads/>, 2019. Accessed 27-11-2020.
- [3] Wikipedia, "Machine learning." https://en.wikipedia.org/wiki/Machine_learning, Dec. 2020. Accessed 27-11-2020.
- [4] A. Wilson, "Supervised learning." <https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590>, Dec. 2020. Accessed 27-11-2020.
- [5] T. Wood, "Unsupervised learning." <https://deepai.org/machine-learning-glossary-and-terms/unsupervised-learning>, Dec. 2020. Accessed 27-11-2020.
- [6] Wikipedia, "Unsupervised learning." https://en.wikipedia.org/wiki/Unsupervised_learning, Dec. 2020. Accessed 27-11-2020.
- [7] Wikipedia, "Cluster analysis." https://en.wikipedia.org/wiki/Cluster_analysis, Dec. 2020. Accessed 27-11-2020.
- [8] Scholarpedia, "Reinforcement learning." http://www.scholarpedia.org/article/Reinforcement_learning, Dec. 2020. Accessed 27-11-2020.