

## Machine Learning Algorithms in Autonomous Cars

2016-11-03

Machine learning algorithms are now used extensively to find solutions to different challenges ranging from financial market predictions to self-driving cars. With the integration of sensor data processing in a centralized electronic control unit (ECU) in a car, it is imperative to increase the use of machine learning to perform new tasks. Potential applications include driving scenario classification or driver condition evaluation via data fusion from different internal and external sensors – such as cameras, radars, lidar or the Internet of Things.

Anshul Saxena, software expert at Visteon's technical center in Karlsruhe, Germany, provides a technical review of the use of machine learning algorithms in autonomous cars, and investigates the reusability of an algorithm for multiple features.

The applications running a car's infotainment system can receive information from sensor data fusion systems and have, for example, the ability to direct the vehicle to a hospital if it senses that something is wrong with the driver. This machine learning-based application can also incorporate the driver's gesture and speech recognition, and language translation. The algorithms can be classified as a supervised algorithm and an unsupervised algorithm. The difference between the two is how they learn.

Supervised algorithms learn using a training dataset, and keep on learning until they reach the desired level of confidence (minimization of probability error). They can be sub-classified into classification, regression and dimension reduction or anomaly detection.

Unsupervised algorithms try to make sense of the available data. That means an algorithm develops a relationship

within the available data set to identify patterns, or divides the data set into subgroups based on the level of similarity between them. Unsupervised algorithms can be largely sub-classified into clustering and association rule learning.

There is now another set of machine learning algorithms called reinforcement algorithms, which fall somewhere between supervised and unsupervised learning. In supervised learning, there is a target label for each training example; in unsupervised learning, there are no labels at all; and reinforcement learning has sparse and time--delayed labels – the future rewards.

Based only on those rewards, the agent has to learn to behave in the environment. The goal in reinforcement learning is to develop efficient learning algorithms, as well as to understand the algorithm's merits and limitations. Reinforcement learning is of great interest because of the large number of practical applications that it can potentially address, ranging from problems in artificial intelligence to operations research or control engineering – all relevant for developing a self-driving car. This can be classified as direct learning and indirect learning.

One of the main tasks of any machine learning algorithm in the self-driving car is continuous rendering of the surrounding environment and the prediction of possible changes to those surroundings. These tasks are mainly divided into four sub-tasks:

- Object detection
- Object Identification or recognition Object classification
- Object localization and prediction of movement

Machine learning algorithms can be loosely divided into four categories: regression algorithms, pattern recognition, cluster algorithms and decision matrix algorithms. One category of machine learning algorithms can be used to execute two or more different subtasks. For example, regression algorithms can be used for object detection as well as for object localization or prediction of movement.

## Regression Algorithms

This type of algorithm is good at predicting events. Regression analysis estimates the relationship between two or more variables, compare the effects of variables measured on different scales and are mostly driven by three metrics, namely:

- The number of independent variables

- The type of dependent variables
- The shape of the regression line.

In ADAS, images (radar or camera) play a very important role in localization and actuation, while the biggest challenge for any algorithm is to develop an image-based model for prediction and feature selection.

Regression algorithms leverage the repeatability of the environment to create a statistical model of the relation between an image and the position of a given object in that image. The statistical model can be learned offline and provides fast online detection by allowing image sampling. Furthermore, it can be extended to other objects without requiring extensive human modeling. As an output to the online stage, the algorithm returns an object position and a confidence on the presence of the object.

These algorithms can also be used for long learning, short prediction. The type of regression algorithms that can be used for self-driving cars are Bayesian regression, neural network regression and decision forest regression, among others.

## Pattern Recognition Algorithms (Classification)

In ADAS, the images obtained through sensors possess all types of environmental data; filtering of the images is required to recognize instances of an object category by ruling out the irrelevant data points. Pattern recognition algorithms are good at ruling out these unusual data points. Recognition of patterns in a data set is an important step before classifying the objects. These types of algorithms can also be defined as data reduction algorithms.

These algorithms help in reducing the data set by detecting object edges and fitting line segments (polylines) and circular arcs to the edges. Line segments are aligned to edges up to a corner, then a new line segment is started. Circular arcs are fit to sequences of line segments that approximate an arc. The image features (line segments and circular arcs) are combined in various ways to form the features that are used for recognizing an object.

The support vector machines (SVM) with histograms of oriented gradients (HOG) and principle component analysis (PCA) are the most common recognition algorithms used in ADAS. The Bayes decision rule and K nearest neighbor (KNN) are also used.

## Clustering

Sometimes the images obtained by the system are not clear and it is difficult to detect and locate objects. It is also possible that the classification algorithms may miss the object and fail to classify and report it to the system. The reason could be low-resolution images, very few data points or discontinuous data. This type of algorithm is good at

discovering structure from data points. Like regression, it describes the class of problem and the class of methods. Clustering methods are typically organized by modeling approaches such as centroid-based and hierarchical. All methods are concerned with using the inherent structures in the data to best organize the data into groups of maximum commonality. The most commonly used type of algorithm is K-means, Multi-class Neural Network.

## Decision Matrix Algorithms

This type of algorithm is good at systematically identifying, analyzing, and rating the performance of relationships between sets of values and information. These algorithms are mainly used for decision making. Whether a car needs to take a left turn or it needs to brake depends on the level of confidence the algorithms have on the classification, recognition and prediction of the next movement of objects. These algorithms are models composed of multiple decision models independently trained and whose predictions are combined in some way to make the overall prediction, while reducing the possibility of errors in decision making. The most commonly used algorithms are gradient boosting (GDM) and AdaBoosting.

As a software expert, Anshul is involved in the development of SmartCore™ and autonomous driving domain controller platforms. He is focused on self-driving car technologies and the effect of the Internet of Things on the auto industry. Anshul is based in Karlsruhe, Germany.