Case Study Notes

Autonomy

Autonomy is the characteristic of a technological system of being independent of human command. The case study underlines the three most important functions that an autonomous system must fulfil are:

1. knowing the car's exact location
2. being able to perceive the car's immediate environment
3. making the correct driving decisions to get from A to B

The challenge highlighted by the case study involves how to implement appropriate technology throughout a town that would support autonomous vehicles.

To make the implementation of autonomous vehicles as safe as possible, the position of the car will need to be known to the nearest centimetre. The case study mentions that this can be achieved through a combination of GPS and high density mapping.

There are advantages associated with employing autonomous vehicles:

1. 81% of crashes are caused by human error; machine errors may be much less due to processing capacities; algorithms to measure stopping distance for example can be executed

However, the issues of implementing autonomous cars will have to be considered:

1. there may be implications for jobs
2. the trolley problem arises at times - decision making of what to do in such a situation may be difficult to achieve
3. the use of neural networks that produce solutions that we don’t really understand
4. beta-testing autonomous cars can be dangerous

Back-propagation

Back-propogation in neural networks computes the exact gradient of some cost function with respect to the weights of the network. The process basically computes some error gradient with respect to the outer layer of the network and propagates those errors towards the input layer, thus allowing you to obtain the exact gradient with respect to all the weights.
In context of autonomous vehicles, back-propagation can be used in convolutional neural networks to achieve minimum errors when distinguishing between objects.

BigO Notation

A mathematical notation to find the time required to run an algorithm as the input increases. = grows linearly as input grows.
The main concern with using Dijkstra's algorithm for autonomous vehicles is that the complexity, or the BigO notation of running the algorithm is where is the number of nodes (major road intersections) between A and B. Such a high complexity makes using this algorithm cumbersome, especially when considering the use of autonomous cars for long distance travel with several road intersections between A and B.

Bounding boxes

Bounding boxes are imaginary boxes that are around objects that are being checked for collision, like pedestrians on or close to the road, other vehicles and signs. There is a 2D coordinate system and a 3D coordinate system that are both being used.

Brute-force

The brute-force method in this context tries to overcome the traveling salesman problem. It involves finding all possible candidates to a solution, and checking individually which of the candidates best fulfils the problem's statement.

In the bus project of this case study, the brute-force method is used to calculate a route that will visit all required locations in the shortest possible time. Although it is the most accurate, it is also impractical since it has a higher complexity than, for example, the nearest-neighbor algorithm.

Convolutional Neural Network (CNN)

CNNs are artificially intelligent structures made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. CNNs take input in the form of images.

In this case study, the input taken is a 32x32x3 pixel input plane. These inputs are taken in by nodes at the input layer. Each pixel has a greyscale value and is part of a 32x32 matrix that is multiplied by a weight matrix, modifying this greyscale matrix, hence the term 'convolution'. ReLU is then used to introduce non-linearity into this convoluted matrix.

Max pooling may also take place to make processing easier by taking the maximum of slices from the 32x32 matrix. For example, from the top-left 2x2 matrix, the average is taken. This has the potential to reduce the dimensions to 16x16.

The advantages of CNNs over ANNs as shown in the case study are:

- the property of shift invariance
- the reduction in the processing required due both to their basic design and to the use of lter strides and pooling
- the reduced memory footprint owing to the use of the same parameters (weights) across each convolution layer.

CNNs are being used in this case study in two manners: object detection and end-to-end learning:

- object detection: classifying objects based on the input matrices received from various sensors
- CNN learns appropriate steering movements by analysing these actions performed by human drivers over a variety of different routes. The given output from the CNN is compared to a dataset of what a real driver would've done, and the errors are reduced by back-propagation.

Cost function

A cost function is a measure of the accuracy with which a neural network performed with respect to its given training sample and the expected output. It also may depend on variables such as weights and biases.

For example, if the expected output is supposed to be 255, but the given output from the CNN is 0, it is said to be very 'costly'. The closer a given output is to the expected output, the less the value of the cost function is.

Deep learning

Deep learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.

Deep learning employs multiple hidden layers in a convolutional neural network in order to gain a deeper understanding of how output varies with input.

Dijkstra's algorithm

Dijkstra's algorithm is an algorithm for finding the shortest paths between nodes in a graph, which may represent, for example, road networks.

It consists of a weighted graph, which has little to do with whether the graph is directed, undirected, or contains cycles. At its core, a weighted graph is a graph whose edges have some sort of value that is associated with them. The value that is attached to an edge is what gives the edge its “weight”.

The major disadvantage of the algorithm is the fact that it does a blind search there by consuming a lot of time waste of necessary resources.

Another disadvantage is that it cannot handle negative edges.

End-to-end learning

End-to-end learning is the idea of outputting complex data types from raw features, for example, audio transcripts, image captures, or steering for self-driving cars. Often has better performance than traditional deep learning. However, it cannot be used for every problem as it requires a lot of labeled data to work optimally.

Feature maps (Activation maps)

The feature map is the output of one filter applied to the previous layer. A given filter is drawn across the entire previous layer, moved one pixel at a time. Each position results in an activation of the neuron and the output is collected in the feature map.

Filter and filter stride

A filter is usually smaller than the input matrix size. A filter can be strode across an input matrix, and by dot products, yields a feature map. Filters are used to convolute the images, by increasing sharpness or highlights or shadows for example. An input matrix of 32x32 can be filtered using a 2x2 filter matrix all the way to a 31x31 filter matrix.

Stride refers to how many pixels are stepped upon each filter. For example, if the 3x3 filter matrix is applied to the top left corner of the 32x32 input matrix, a stride of 1 means that the filter jumps from (0,0) to (1,0) to (2,0) all the way to (29,0), and then to (0,1). A stride of 2 means that the filter jumps from (0,0) to (2,0) to (4,0) all the way to (28,0), and then to (0,2).

Greedy algorithm

A greedy algorithm, as the name suggests, always makes the choice that seems to be the best at that moment. This means that it makes a locally-optimal choice in the hope that this choice will lead to a globally-optimal solution.
It is much cheaper than brute-force or exhaustive search, and taking the next best available choice is usually easy if all the weights are sorted. However, sorting these weights may add to the expense of greedy algorithms. Greedy algorithms require optimal local choices.

Machine learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.
Machine learning is of two types: supervised (based on datasets with some primary expectation of output), and unsupervised (studies how systems can infer a function to describe a hidden structure from unlabeled data)

Max-pooling / Pooling

Max pooling is a way of reducing the dimensions of a feature map in order to make convolutions and output predictions less expensive in terms of system resources. Max pooling a 32x32 matrix, for example, can involve taking the greatest pixel value from 2x2 slices across the matrix.

Multi-layer perceptron (MLP)

It is a type of artificial neural network. An MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable.

Nearest neighbour algorithm

An algorithm to solve the travelling salesman problem. Finds the closest neighbour from point X, and visits it, and then when the destination is reached, returns to point X, and try the second best neighbour. If its used to give a solution to the travelling salesman problem, it will give a short tour, but won't give a nice optimised solution.

The advantages of NN algorithm includes:

- Simple to implement
- Flexible to feature / distance choices
- Naturally handles multi-class cases
- Can do well in practice with enough representative data

The disadvantages of NN algorithm includes:

- Large search problem to find nearest neighbours
- Storage of data
- Must know we have a meaningful distance function

Overfitting

Overfitting is when you train an algorithm too much, making it behave quite strangely. When you fit an algorithm, it tries to model the data provided; over time, if an algorithm is overfitted, the error will increase if another set of data is given. When you introduce unique data to the algorithm, it has the tendency to overfit what it perceive to what it's been taught, especially in CNNs.

Point clouds

Point clouds are large data sets composed of 3D point data. LIDAR scanners are the most common instruments used to collect geographic point cloud data.

Point clouds using LIDAR is useful for autonomous car systems because it allows them to capture incredibly detailed information about a physical object.

Receptive field

A receptive field is the group of pixels from the previous layer's pixel matrix which is what is represented by a neurone or node on the current layer. For example, a neurone in a hidden layer can represent any 5x5 region of the 32x32 hidden matrix. Hence, this neurone is said to have a receptive field of region 5x5 pixels.

Sensor Fusion

Software that combines data from several different sensors to improve performance in a program. By combining the data, the deficiencies of each sensors are cancelled out.

The sensor fusion in this case study is evident: the vehicle uses LIDAR, long-range radar, cameras, and short/medium-range radars. All of these are fused in order to specialise in some aspect and yield the best autonomy for the car.

LIDAR (light detection and ranging)

How it works: laser signals are emitted, laser signals reach an obstacle, signal reflects from the obstacle, signal returns to the receiver, and then a laser pulse is registered.

Advantages: short wavelength lets us detect small objects; a LIDAR can build an exact 3D monochromatic image of an object.

Disadvantages: limited usage in nighttime/cloudy weather, operating altitude is only 500-2000m, quite an expensive technology.

RADAR (radio detection and ranging)

Advantages: RADAR can easily operate in cloudy weather conditions, and at night, longer operating distance.

Disadvantages: shorter wavelength does not allow the detection of small objects; RADAR cannot provide the user with the precise image of an object because of the longer wavelength

Society of Automotive Engineers

SAE International is a global association of more than 128,000 engineers and related technical experts in the aerospace, automotive and commercial-vehicle industries.

Shift invariance (Spatial invariance)

An invariant is a value that can be assured that it will stay true while the program runs. In an invariant system, If is the response of the system to , then is the response of the system to . In a shift invariant system, the would respond to the shift of space, while in a time invariant system, time would be responsible of . This is to say that over a change in ,

VTV and VTI protocols

VTV protocols address technology designed for vehicles to communicate with each other. The communication is wireless, many of these protocols rely on simple comm protocols, which can be used to detect the car near your car to send out a signal. If all cars will be connected to each other using this, we might not need street lights for cars.

A VTI protocol is any communication model that allows the vehicle to share information with infrastructure meant for road control (traffic lights, lane markers, streetlights, signage and parking meters). It is wireless and bidirectional. Can be used to give the user information about parking availability, accidents, roadblocks, traffic, etc.

The challenges faced with VTV and VTI protocols include:

- Ensuring that possible sharing with other wireless users of the radio-frequency spectrum used by V2I communications will not adversely affect V2I technologies' performance;
- Addressing states and local agencies' lack of resources to deploy and maintain V2I technologies;
- Developing technical standards to ensure interoperability;
- Developing and managing data security and addressing public perceptions related to privacy;
- Ensuring that drivers respond appropriately to V2I warnings; and
- Addressing the uncertainties related to potential liability issues posed by V2I.