CIHAN UNIVERSITY-ERBIL

HYBRID DEEP LEARNING MODELS FOR SMART CITIES DATA

Asst. Prof. Dr. Reem Jafar Ismail

Department of Computer Science

SMART CITY

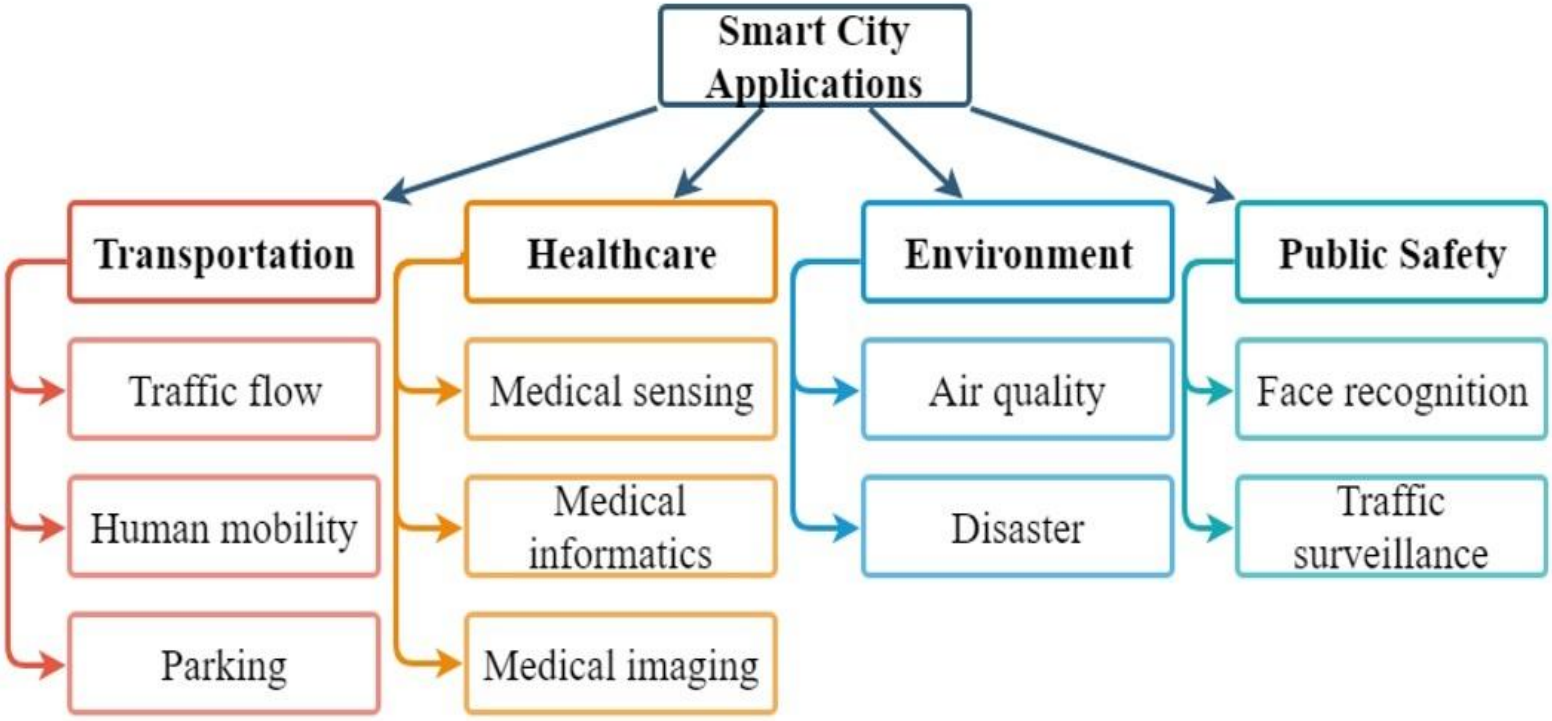
A ‘smart city’ uses digital technologies to provide city services and infrastructure such as transport, energy, health care, water and waste management.

Information and Communications Technologies (ICT) play an important role in connecting data from its resources, securely managing the massive amounts of data generated, and providing the relevant services that are required.

LAYERS OF SMARTNESS

To transform a city infrastructure in “smart”, the city manager should develop three layers of “smartness” as follows:

- The technology base include networks of sensors and connected devices which gather data
- The second layer consists of the deployment of smart applications that process the raw data, translating it into alerts, insight, and actions.
- The third layer involves widely adopting the system by the residents.



SMART CITY APPLICATIONS

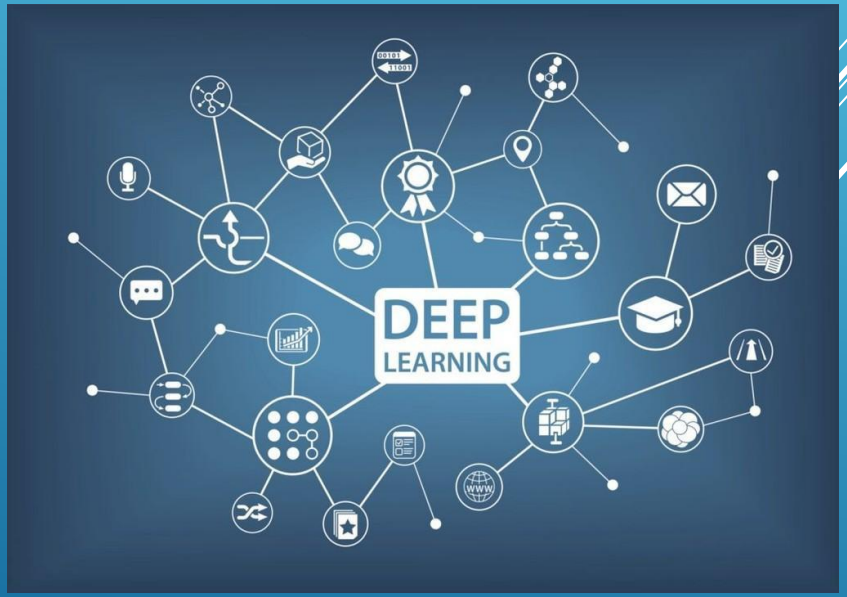
MACHINE LEARNING

- 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.
- The process of learning begins with observations of data in order to look for patterns and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

DEEP LEARNING

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the

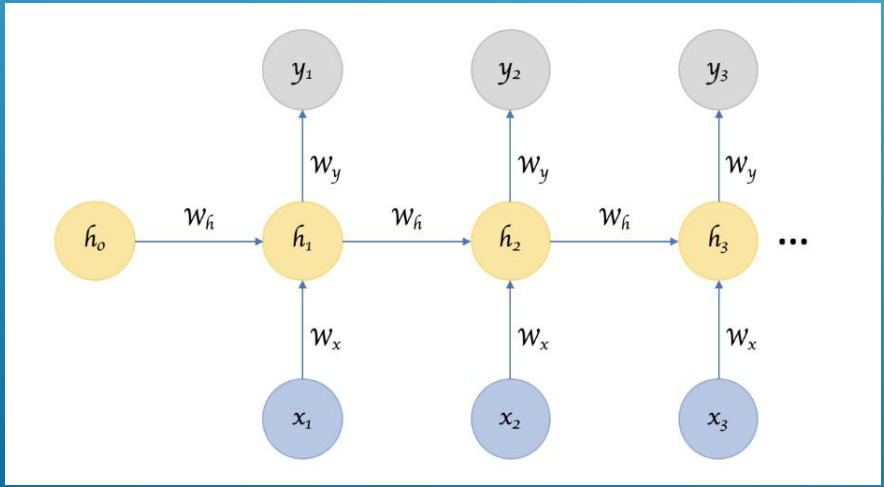
brain called **artificial neural networks**.



ARTIFICIAL NEURAL NETWORK (ANN)

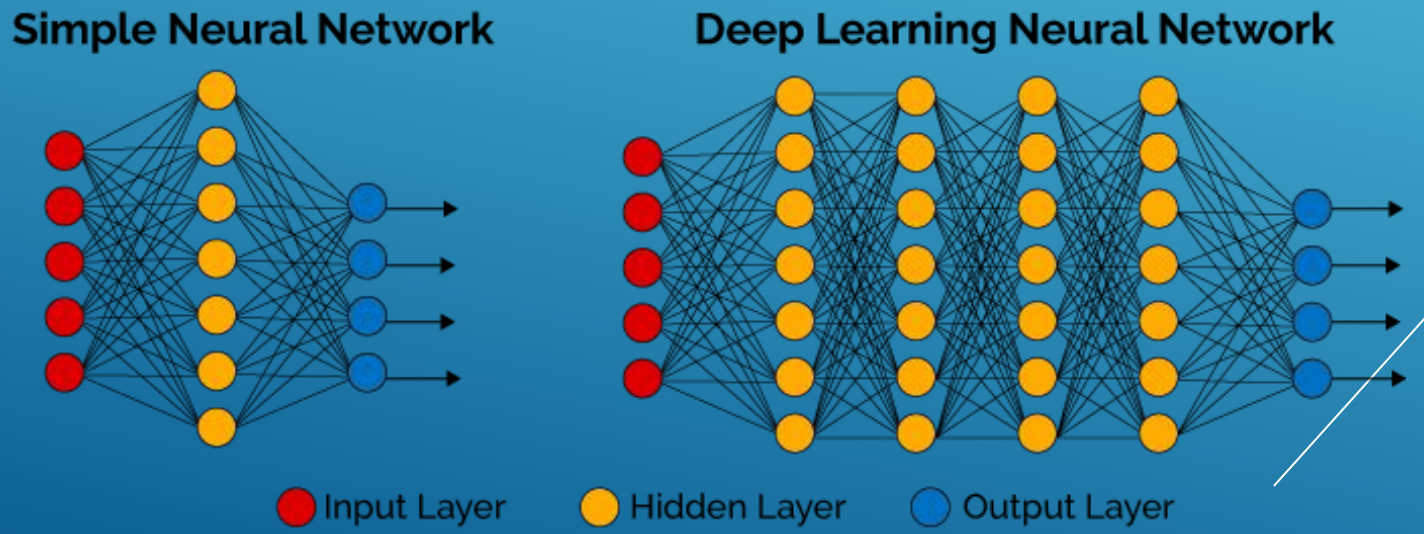
- Neural networks, a biologically-inspired programming concept which enables a computer to learn from observational data.

- An ANN consists of a collection of connected artificial neurons. It can be characterized by the weights between layers of neurons whose output is computed based on some non-linear transformation function. At each layer, neurons compute a weighted sum of the inputs from the previous layer, using Wx + b where W is a weight vector and b is a bias, and pass the result through a non-linear activation function .



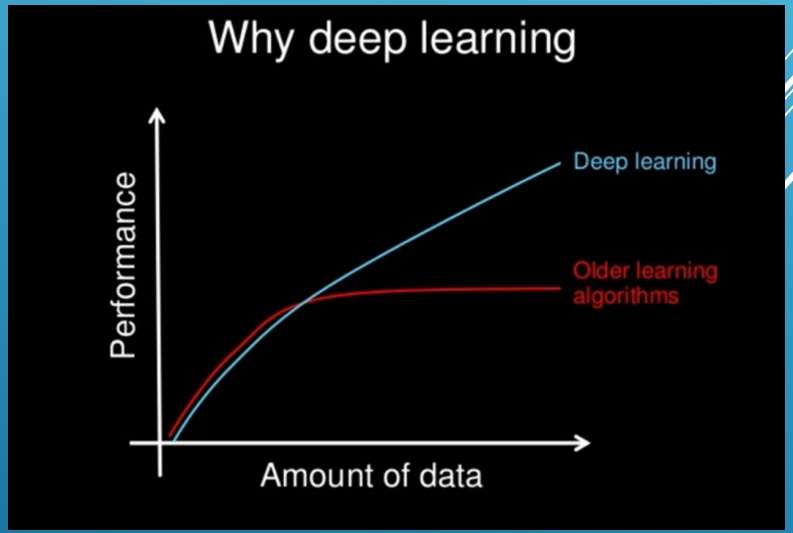
- These artificial networks may be used for predictive modeling, adaptive control and applications where they can be trained via a dataset. Self-learning resulting from experience can occur within networks, which can derive conclusions from a complex or unrelated set of information.

We add depth by adding more hidden **states** and **layers**, where additional hidden layers added between input to hidden layer and between output to hidden layer or by increasing the number of states in the hidden layer itself.



WHEN TO USE DEEP LEARNING?

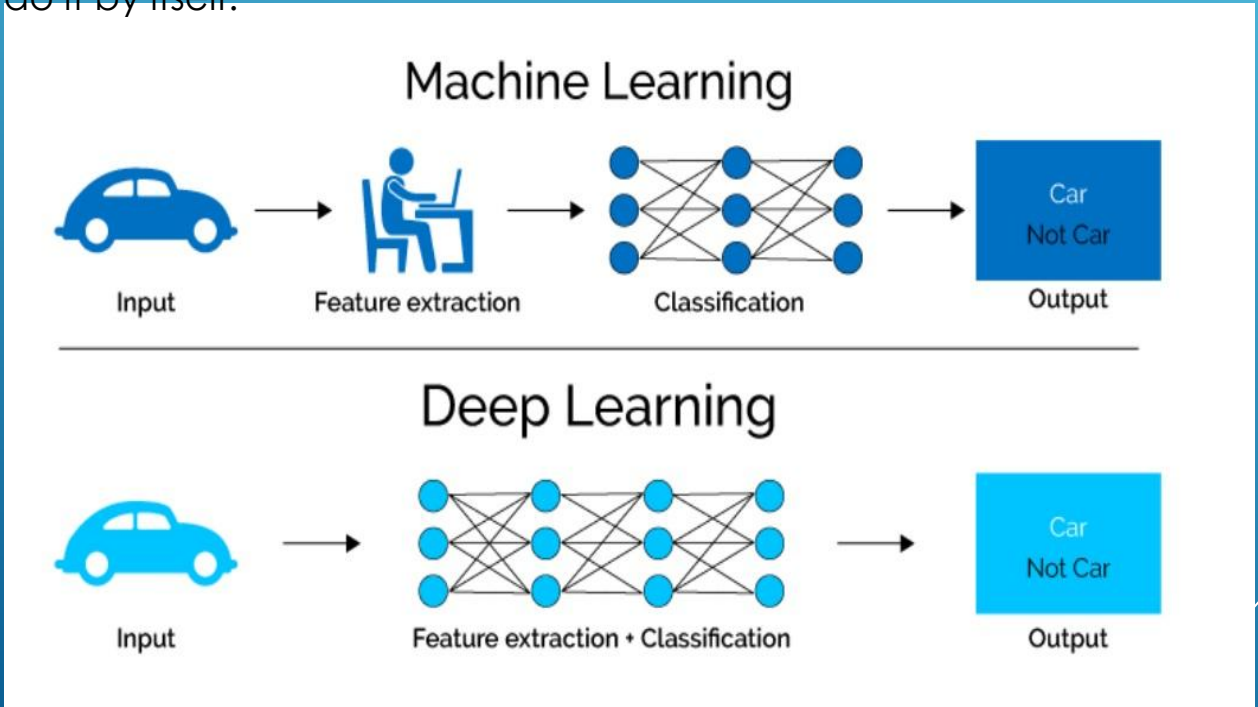
- **Deep Learning** out perform other techniques if the ***data size is large***.
- **Deep Learning** techniques need to have ***high end infrastructure*** to train in reasonable time.
- **Deep Learning** used when it comes to *complex problems* ***such as image classification, natural language processing, and speech recognition***.



Deep Learning used when there is no expert or observatory for feature

extraction of the information, like in smart cities where the systems should

do it by itself.



DATA REPRESENTATION

TEXT DATA REPRESENTATION

- Processing textual data heavily relies on natural language processing techniques. In deep learning, textual data can be represented as temporal sequence of variable lengths, similar to sensor data representation, i.e., x = {x1,x2,...xt,...,xT}, where x denotes a sentence or text block and xt a vector representing a word at time step t.

- One of the typical applications is to analyze short textual messages posted on social media platforms (e.g., Twitter) to discover knowledge from multiple perspectives in real time, e.g., social event detection, traffic event detection, disease activity tracking, and natural disease monitoring.

IMAGE AND VIDEO DATA REPRESENTATION

- A digital image is usually represented as a three-dimensional matrix with numeric values. Each point in an image is called a pixel, denoting the color value at that specific position. The number of rows and columns of an image is fixed and indicates the resolution of the image.

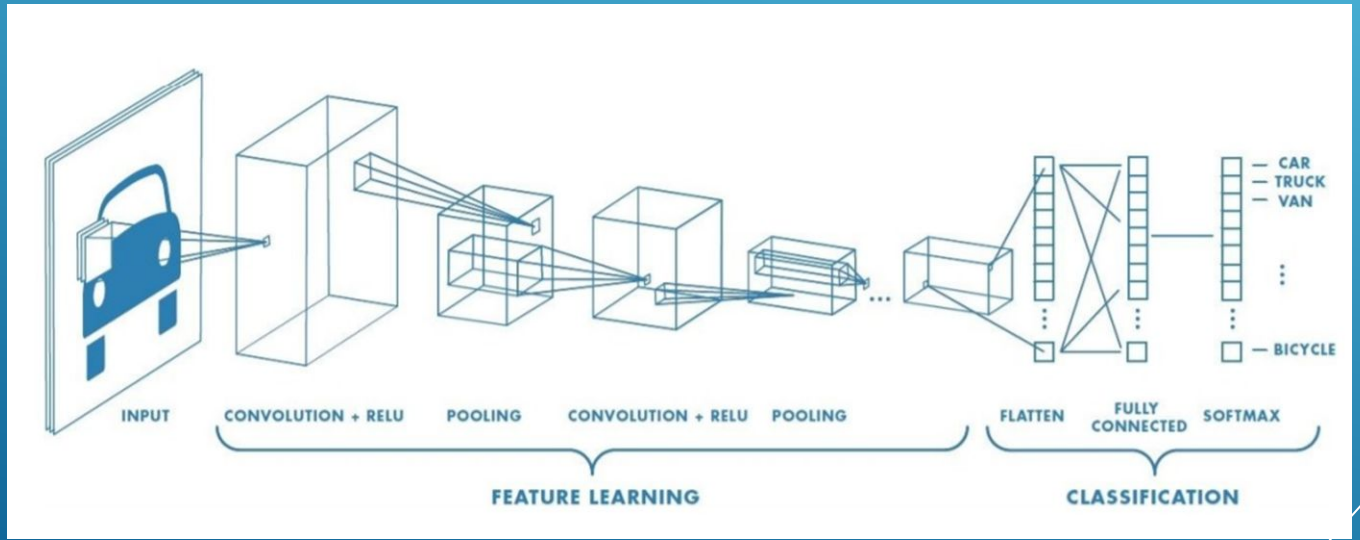
- Image and video data is important in smart cities and can be collected in many ways using: **cameras** which are widely installed in city areas for public safety and criminal investigation with remotely controllable devices like (i.e., sensors, cameras, and actuators); **satellite images** are used to detect and classify objects on earth for environment protection and weather forecasting; citizens use **smartphones** to take images and videos to record daily lives and report social events.

DEEP LEARNING MODELS

- Convolutional Neural Networks (CNN)
- Recurrent Neural Networks (RNN)

CONVOLUTIONAL NEURAL NETWORKS (CNN)

- Convolutional neural networks are usually composed by a set of layers that can be grouped by their functionalities.



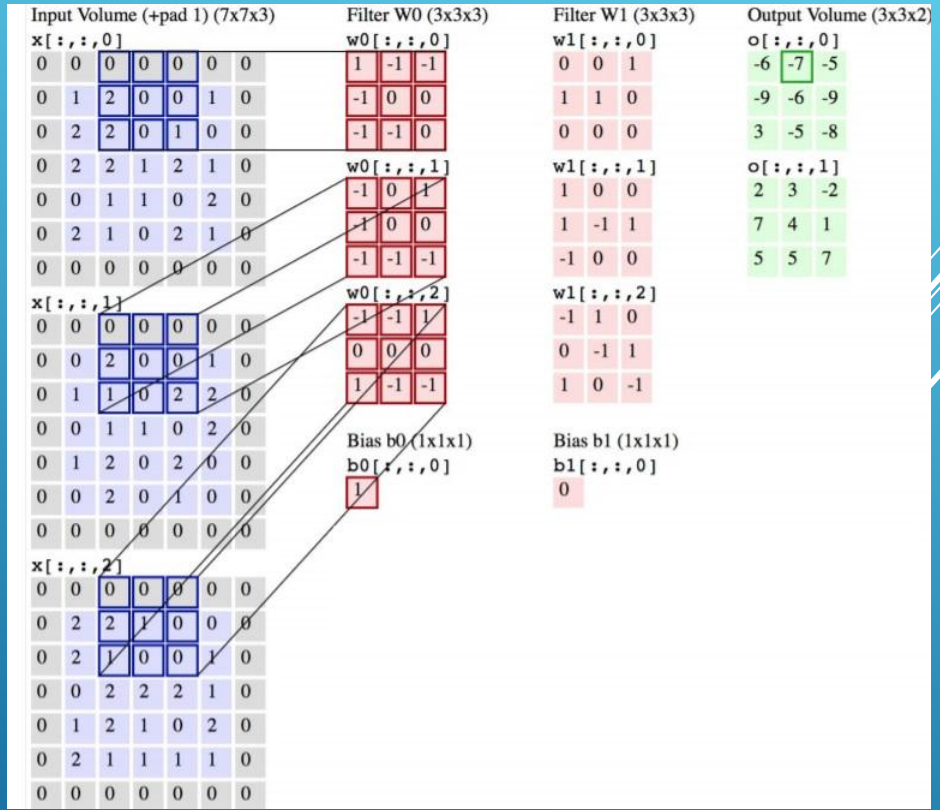
Convolutional neural networks use three basic ideas:

- Convolution layer – feature extraction
- Pooling layer - feature extraction
- Fully connected layer - classification

CONVOLUTION LAYER

The process is a 2D convolution on the inputs

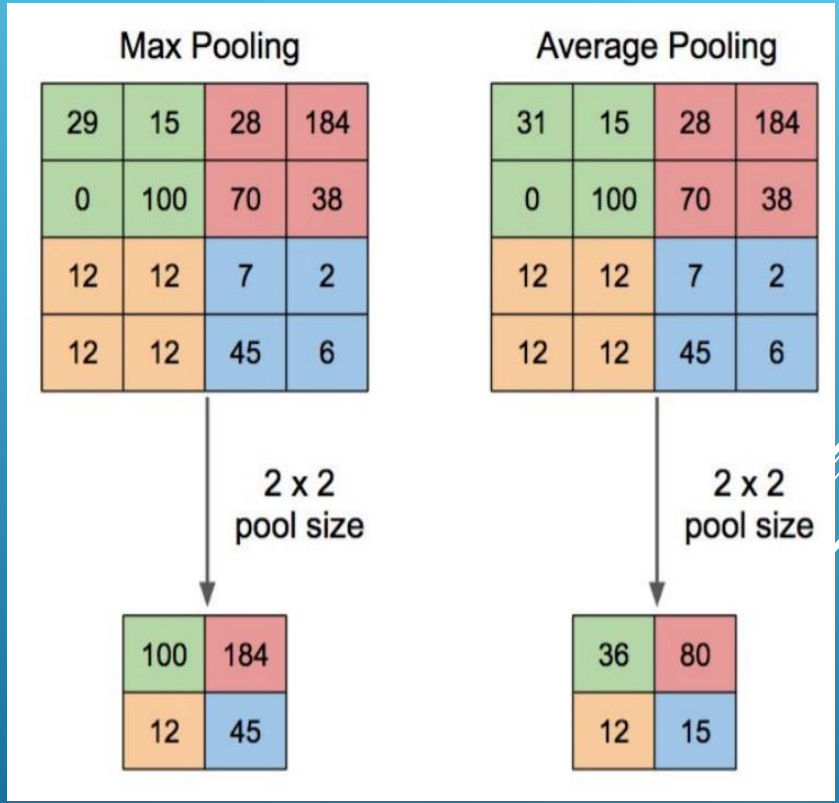
Using the “dot products” between weights and inputs are “integrated”.



POOLING LAYER

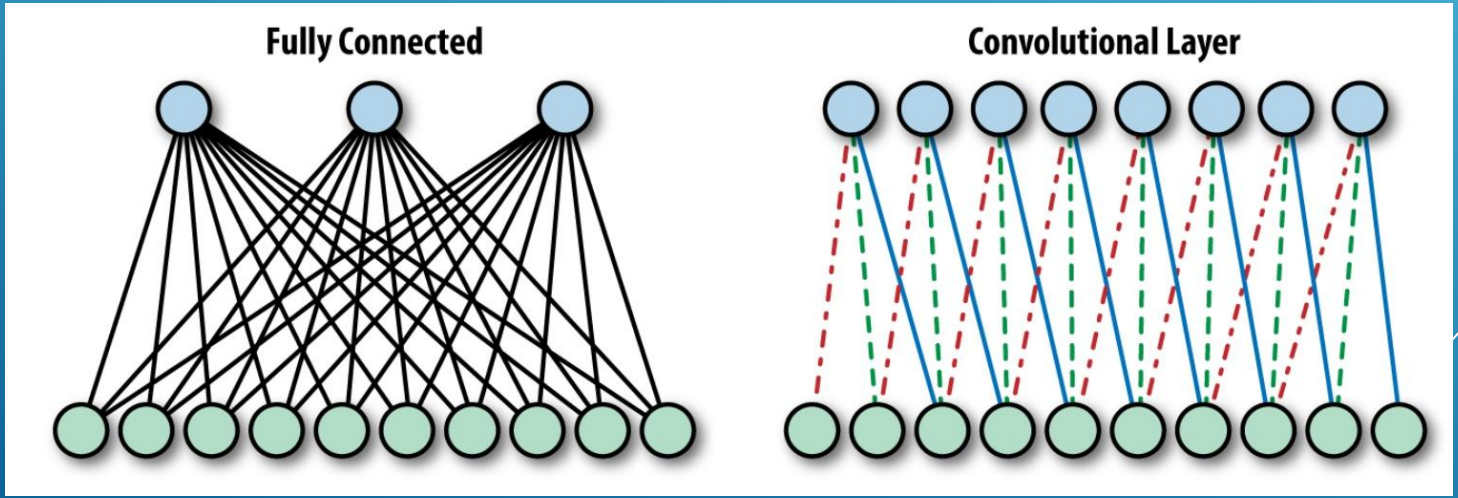
Pooling is an aggressive filtering process which uses smaller filter size to discard information.

-Max pooling
-Average pooling

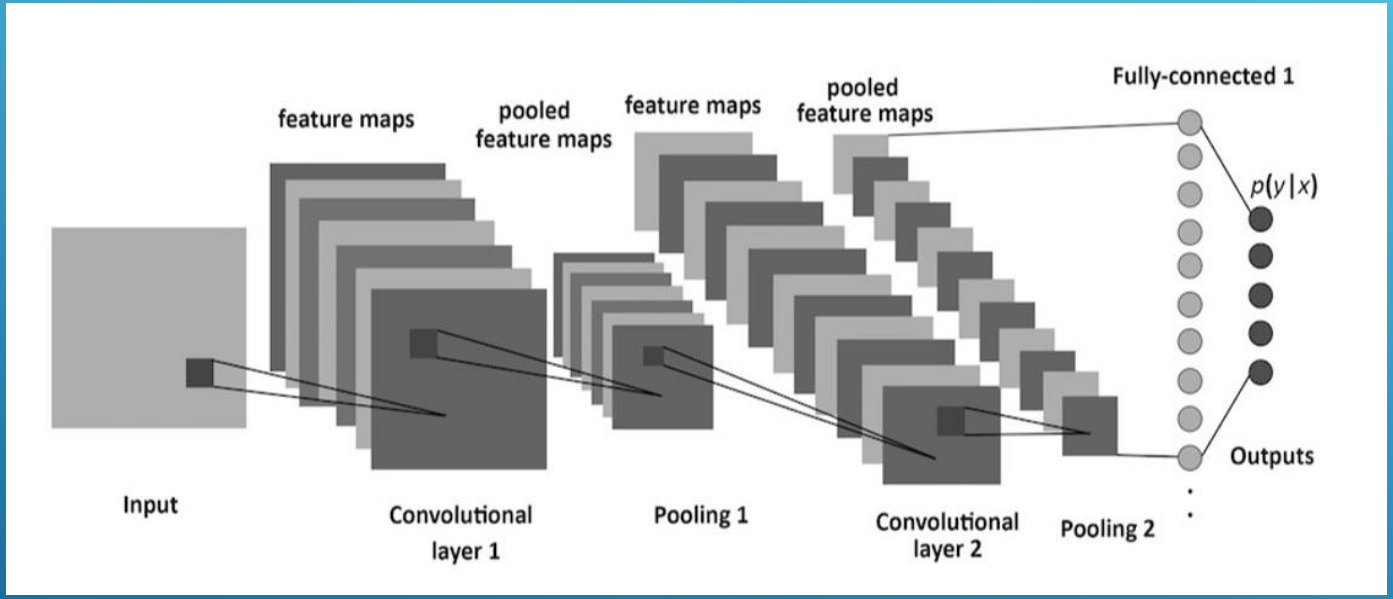


FULLY CONNECTED LAYERS

- In a fully connected layer (left), each unit is connected to all units of the previous layers. In a convolutional layer (right), each unit is connected to a constant number of units in a local region of the previous layer.

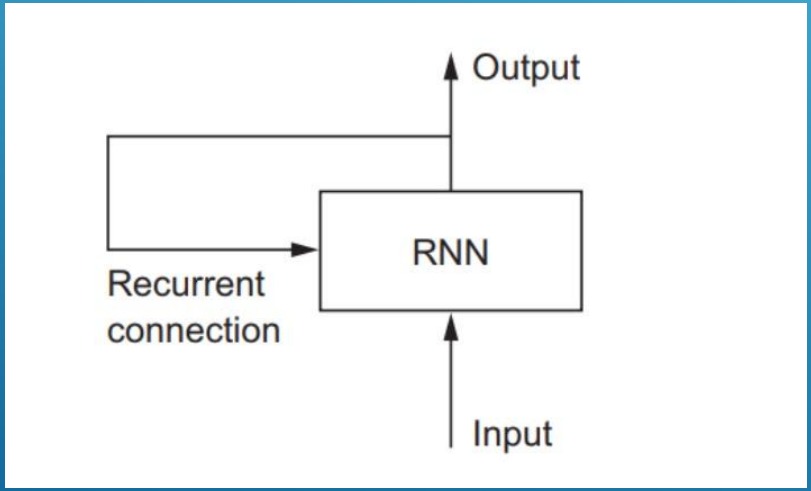


CONVOLUTIONAL NEURAL NETWORKS STAGES

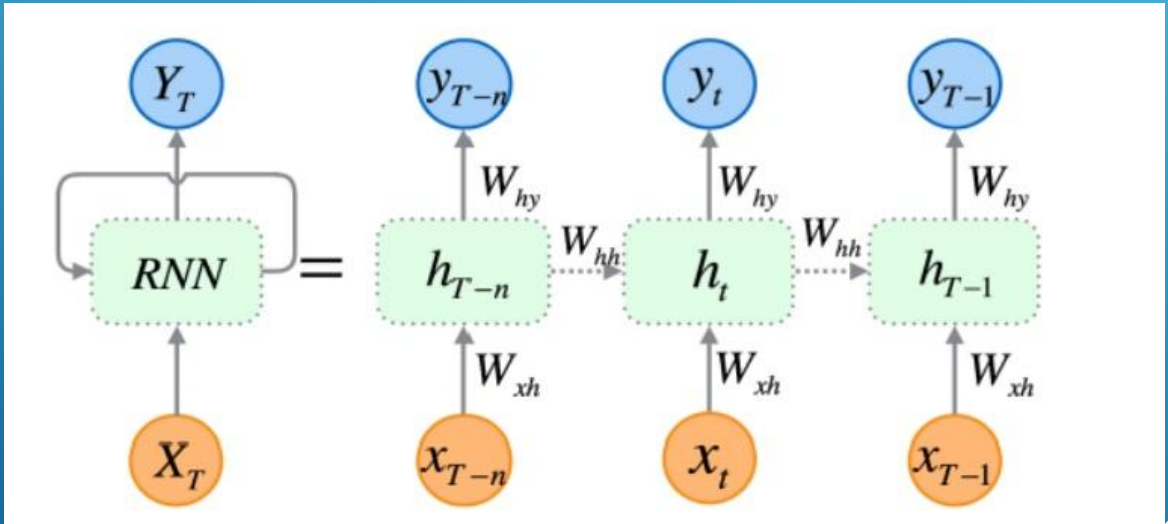


RECURRENT NEURAL NETWORKS (RNN)

- *Recurrent* means the output at the current time step becomes the input to the next time step. At each element of the sequence, the model considers not just the current input, but what it remembers about the preceding elements.



- RNN process data that can be modelled as temporal sequences of variable lengths, x = (x1,...,xT). At each time step t, the hidden state ht of the RNN is updated using ht = f(ht−1,xt), where f is a non-linear function, its function as complex as a **long short-term memory** unit. RNNs use the internal states to capture dependency among input data in a sequence, which makes them suitable to tasks such as natural language processing, speech recognition.



OVERVIEW OF DEEP LEARNING MODELS FOR DIFFERENT TYPE OF CITY DATA

Model	Input representation	City data
RNN	Time sequence data $\{x_1, x_2, ..., x_t\}$	traffic flow or speed (sensor) air quality data (sensor) healthcare data (sensor) human activity (sensor)
CNN	2D or 3D matrix $\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,n} \end{bmatrix}$	human mobility (sensor) human activity (sensor) parking lot image medical image disaster image twitter data (text) face image vehicle image

HYBRID MODEL

- Hybrid model is referred to as the combination of more than one type of deep learning models. Recent research has seen a number of such hybrid models applied in different application domains in smart cities.
- Hybrid models that combine CNN with RNN work well and tend to perform better than any single models.

SOME HYBRID NEURAL NETWORKS MODELS FOR SMART CITY APPLICATIONS

Contribution	Application	Data/ Input	Hybrid Neural Network Models
End-to-end travel time estimation	Transportation/ Human mobility	Vector/ Distance of path Weather condition Driver habit Day of week	CNN: to predict the citywide crowd flow RNN: to estimate the local path
license plates recognition	Public safety/ Traffic Monitoring	Image/ Segmented license plates	CNN: for feature extraction RNN: for recognition
Detection and classification of dynamic hand gestures	Healthcare/ Medical sensing	Video/ Color Optical flow Depth	CNN: enable gesture classification based on hand or figure moving types RNN: for recognition

DO YOU WANT TO LIVE IN GREEN OR SMART CITY?