INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

# Impact on Blockchain-based AI/ML-enabled Big data analytics for Cognitive Internet of Things environment

**Dr. Ashok Kumar Das**

**IEEE Senior Member**
**Associate Professor**
Center for Security, Theory and Algorithmic Research
International Institute of Information Technology, Hyderabad

E-mail: *ashok.das@iiit.ac.in*
URL: `http://www.iiit.ac.in/people/faculty/ashokkdas`
`https://sites.google.com/view/iitkgpakdas/`

# AI-enabled Blockchain-based Big data analytics in Cognitive Internet of Things (CIoT)

**Reference:** Ankush Mitra, Basudeb Bera, Ashok Kumar Das, Sajjad Shaukat Jamal, and Ilsun You. "Impact on Blockchain-based AI/ML-enabled Big data analytics for Cognitive Internet of Things environment," in ***Computer Communications (Elsevier)***, 2022. (2021 SCI Impact Factor: 5.047)

# Cognitive Internet of Things (CIoT)



- Internet of Things (IoT) has become an emerging technology due to a huge enhancement of Information and Communications Technology (ICT). IoT compromises a large number of smart devices, called IoT devices, which can be either physical or virtual objects.
- Cognitive Internet of Things (CIoT) has now become a new network model which is enhancement of IoT. Likewise IoT network, physical and virtual objects are part of CIoT, which work with minimum human intervention and they also communicate with each other based on a “context-aware perception-action cycle”.
- CIoT is thus considered as a field of science where IoT and cognitive computing are applied to make the IoT systems smarter. It provides some kind of thinking ability to the IoT systems.

# Cognitive Internet of Things (CIoT)

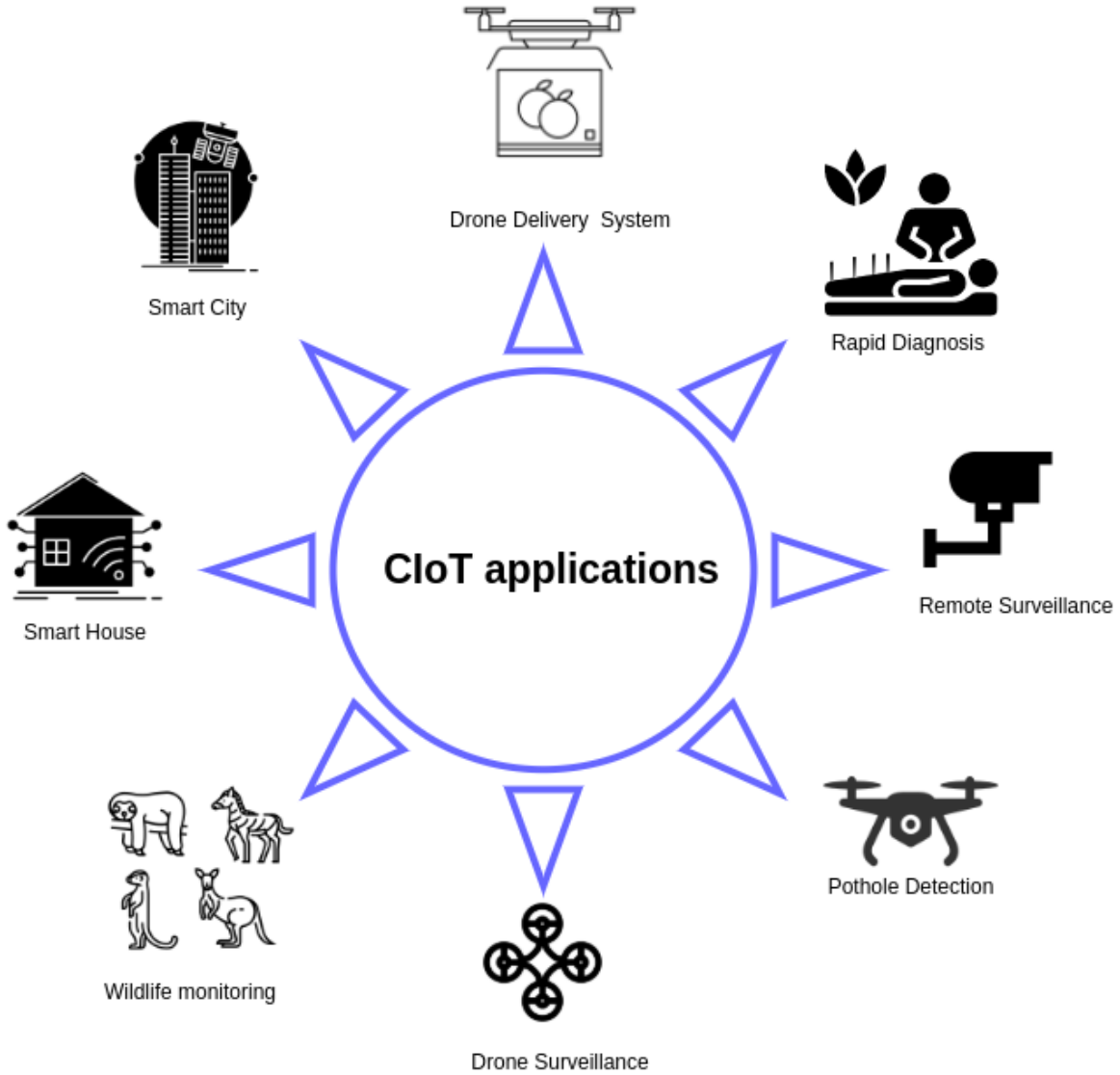



Figure: Various CIoT-enabled applications

# Attacks related to AI/ML

AI/ML security becomes an emerging topic in the computer science field in order to make correct and accurate predictions on non-poisonous (corrupted) data. For instance, typically a huge volume of data generated by the IoT smart devices in CIoT can be stored in cloud server(s). As the cloud servers are semi-trusted, there is a possibility by the insider attackers of the cloud servers to perform several attacks.

- *Adversarial input attacks*: These are specially crafted on inputs that have been developed with the aim of being reliably mis-classified in order to evade detection.
- *Data poisoning attacks*: The attacker can then insert false data, alter the label of the data, and also remove the data or insert random noise to the data to poison the training data.
- *Model attacks*: In such type of attack, an attacker may pollute the model's hyper-parameters (that is, the parameters that are learned using AI/ML).
- *Model stealing attacks*: Such kinds of attack scenarios are used to steal or duplicate models or recover training data membership.

# Convolution Neural Networks (CNN)



Convolution Neural Networks (CNN) is a deep neural network that is generally used in computer vision.
CNN mainly contains three types of layers:

- *Convolution layer*: It is a convolution tool that splits various features of the input images for analysis.
- *Pooling layer*: The main goal of this layer is that it decreases the size of the convoluted feature map in order to reduce computation costs.
- *Fully connected layer*: This layer applies the output of the convolution layer for making prediction about the best description for the inputs (images).

# Convolution Neural Networks (CNN)

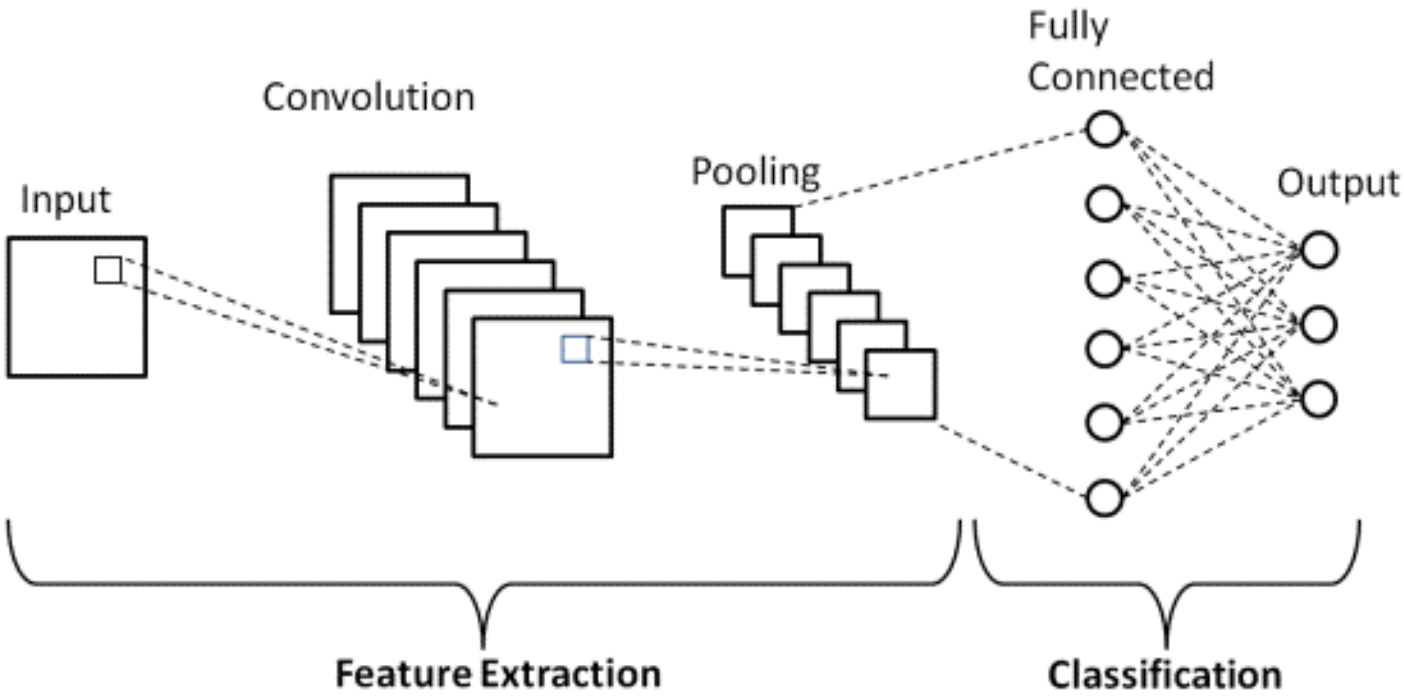
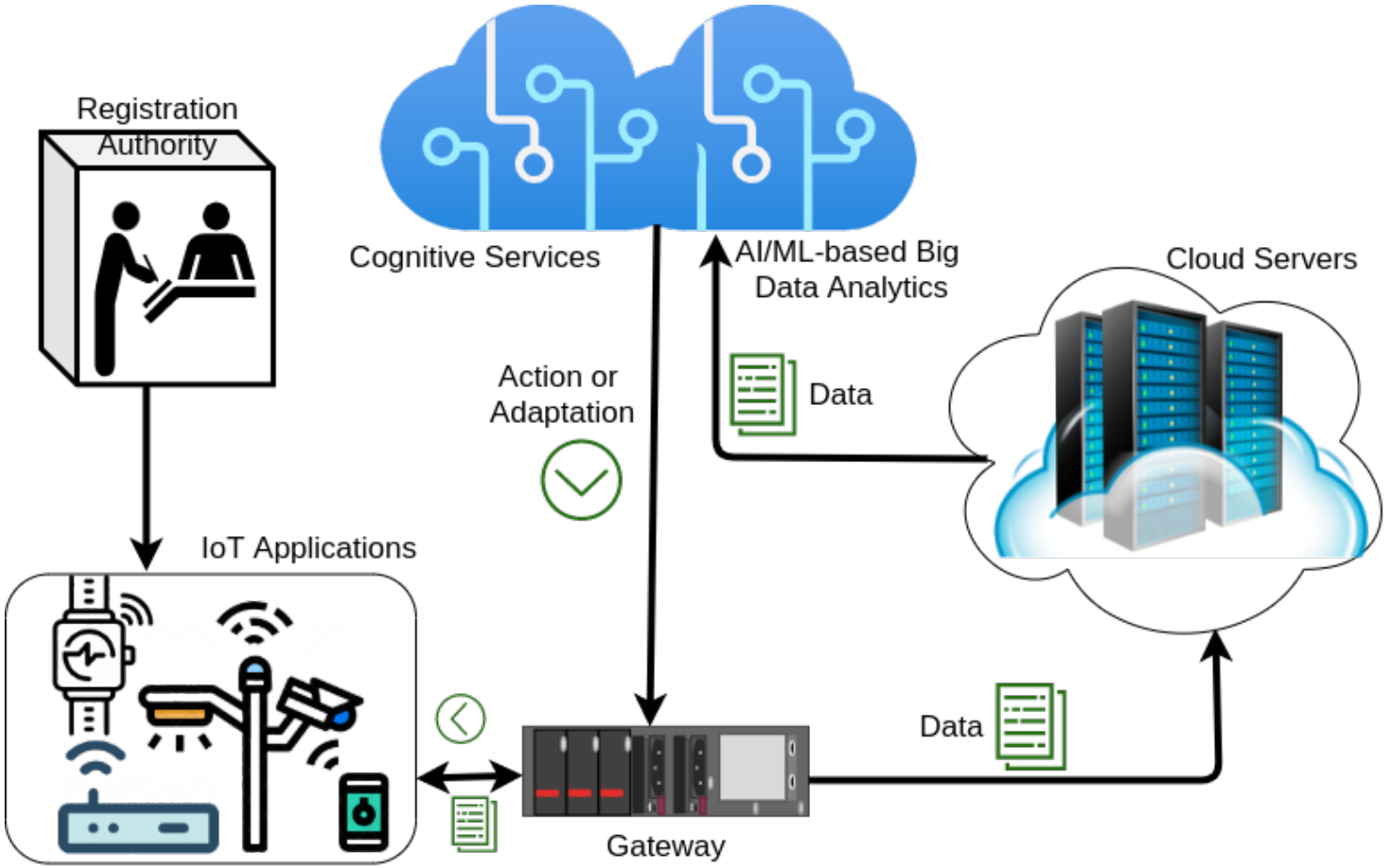
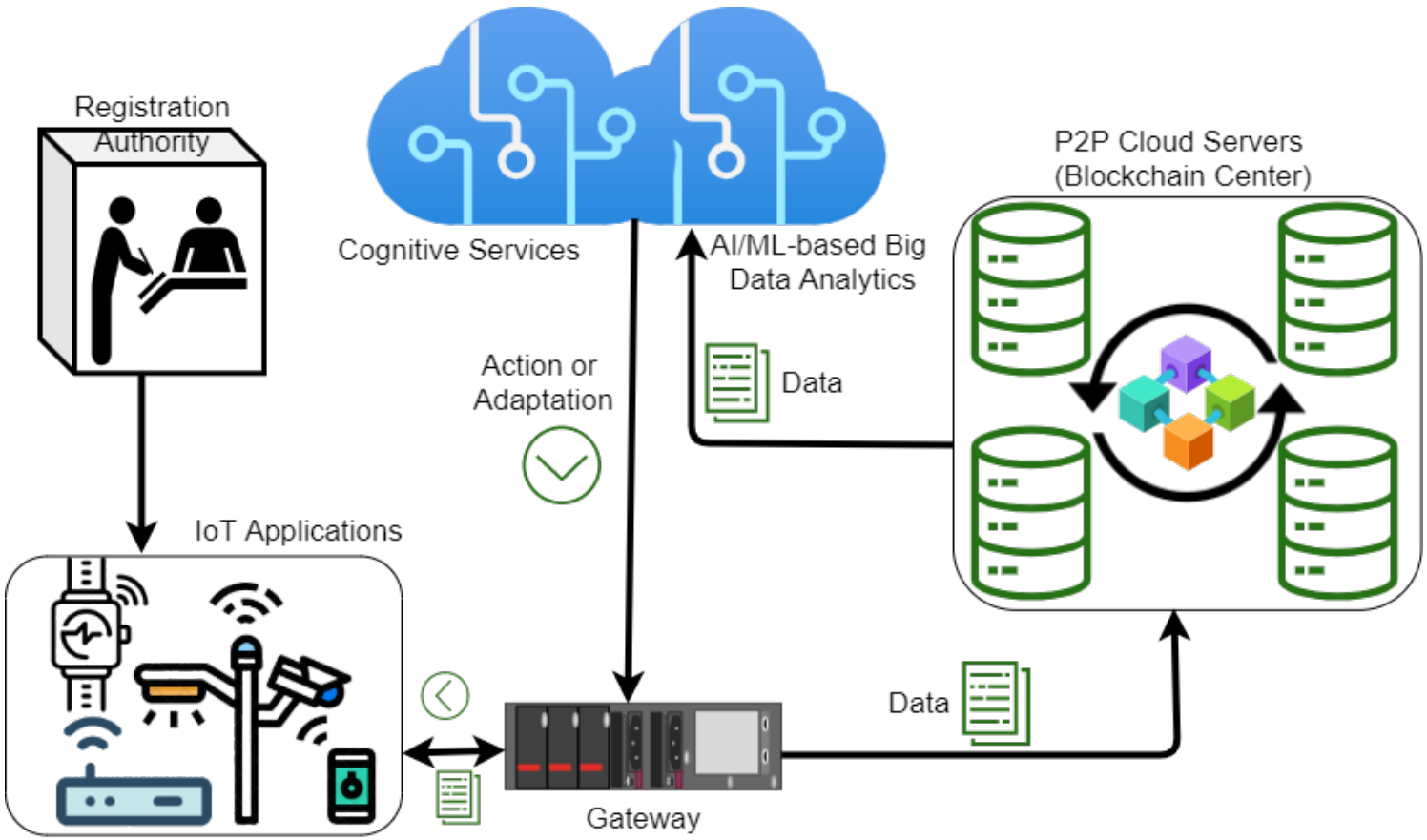



Figure: CNN architecture

# CIoT network model without blockchain



# Blockchain-envisioned CIoT network model

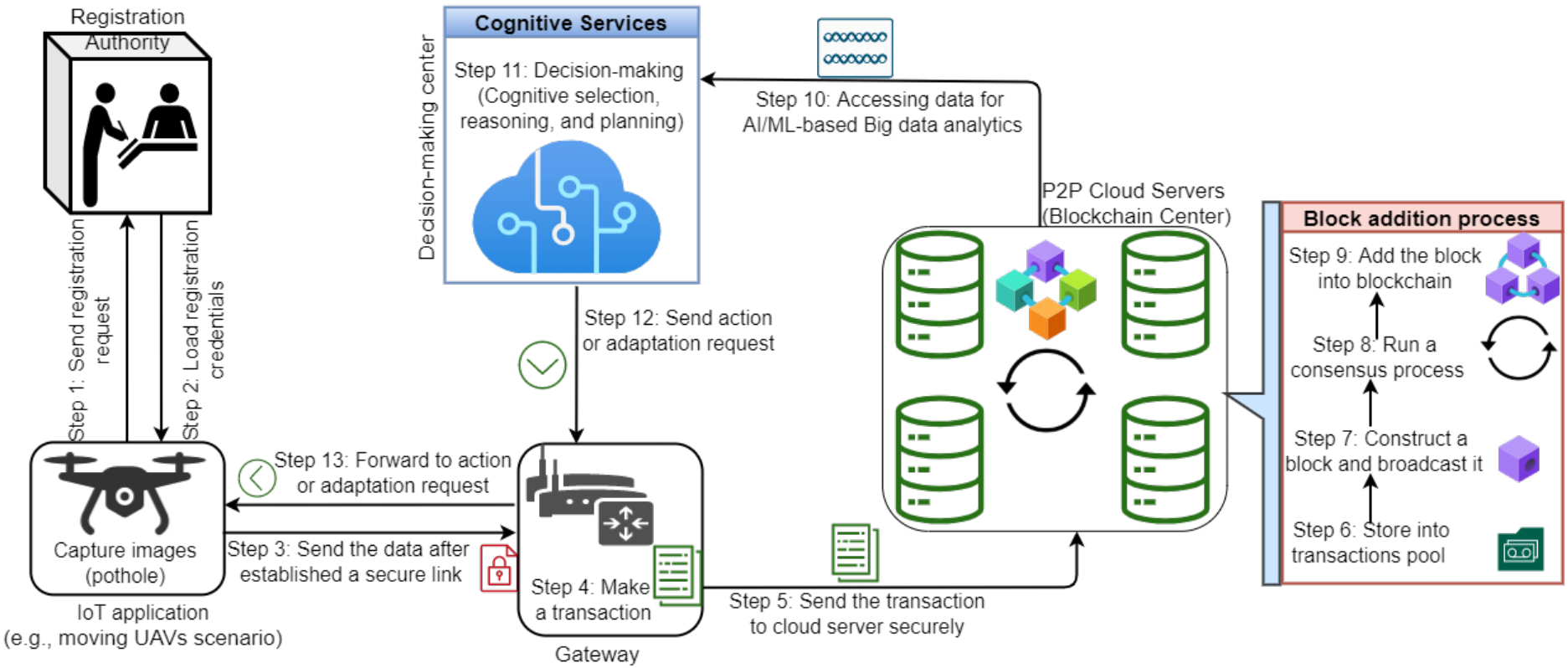


# AI-enabled Blockchain-based Big data analytics in CIoT

| **Block Header** |
| --- |
| Block Version |
| Previous Block Hash (PBH) |
| Merkle Tree Root (MTR) |
| Timestamp |
| Signer's Public Key |
| **Block Payload (Transactions)** |
| (Transactions ($Tx$), ECDSA.Sig($Tx$)) |
| Current Block Hash (CBH) |

Figure: A block structure in public blockchain

# Overall process in blockchain-based AI/ML-enabled Big data analytics



# Experimental Results

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

## Data Sets:

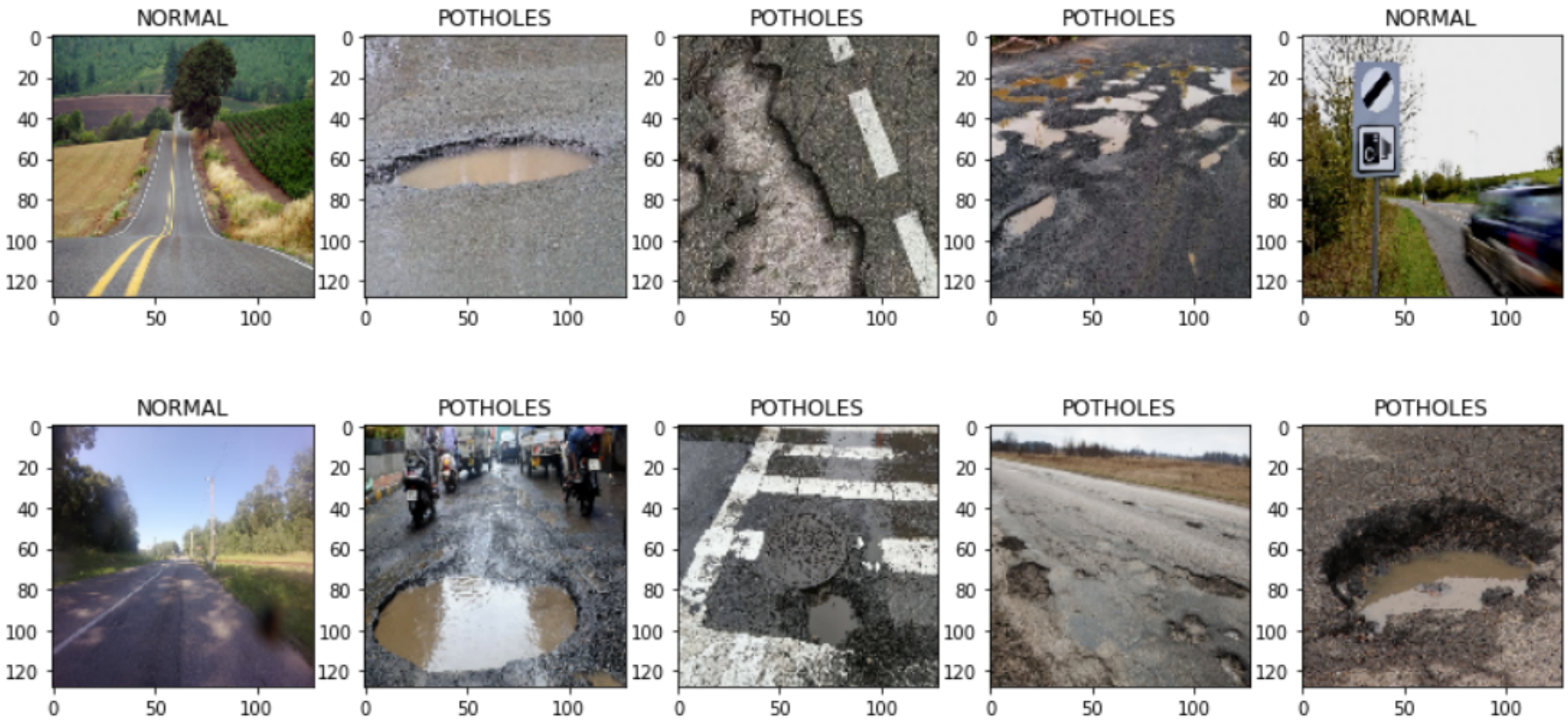
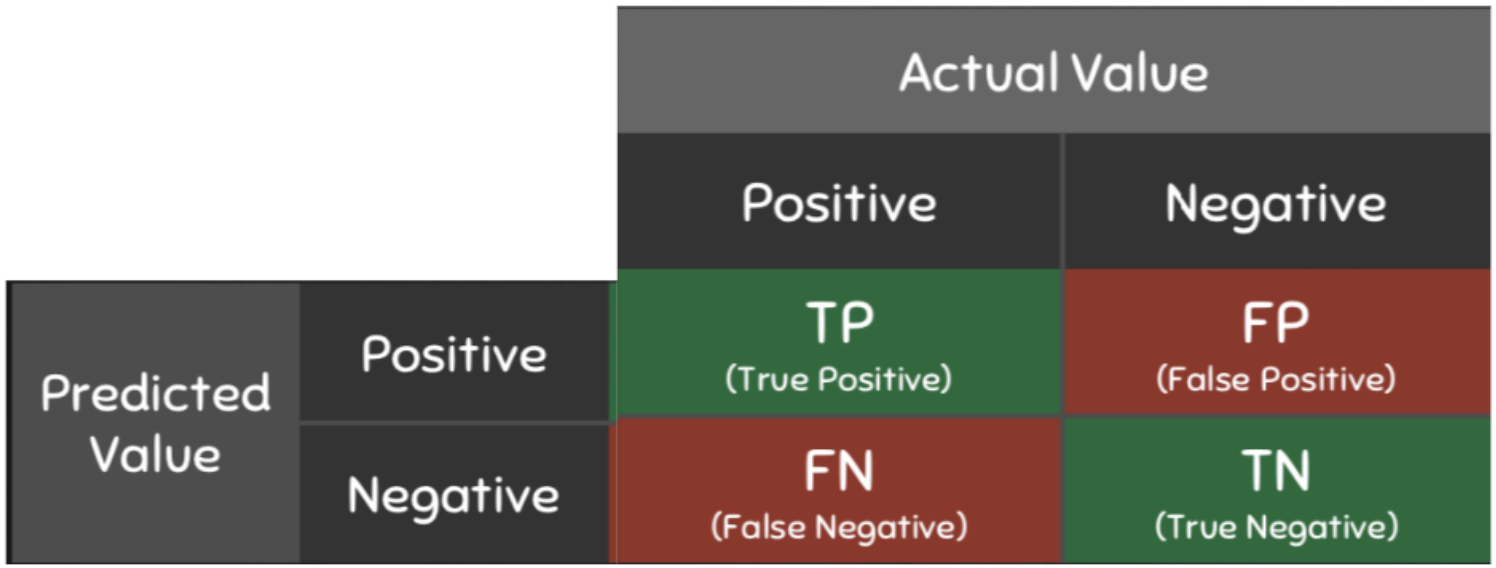



Figure: Pothole data sets (Source: `https://www.kaggle.com/atulyakumar98/pothole-detection-dataset`)

# Metrics Used in Experiment




- True Positive (TP) : Observation is positive, and is predicted to be positive.
- False Negative (FN) : Observation is positive, but is predicted negative.
- True Negative (TN) : Observation is negative, and is predicted to be negative.
- False Positive (FP) : Observation is negative, but is predicted positive.

Figure: Structure of a confusion matrix

# Metrics Used in Experiment



- *Accuracy*:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- *Recall*:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- *Precision*:

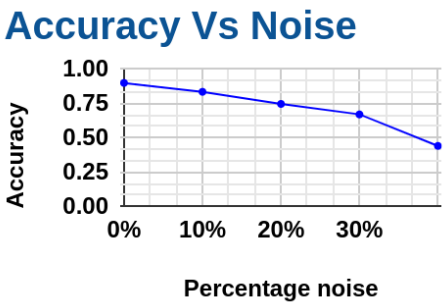
$$\text{Precision} = \frac{TP}{TP + FP}$$

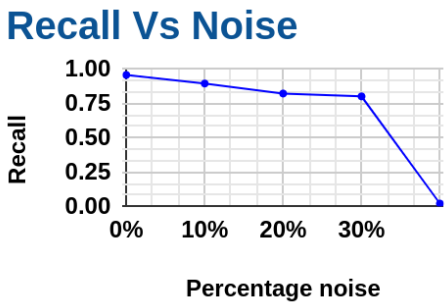
- *F1 score*:

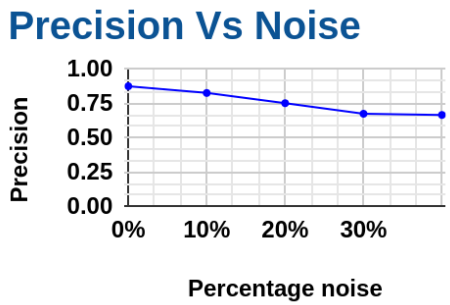
$$\text{F1 score} = \frac{2 * TP}{2 * TP + FP + FN}$$

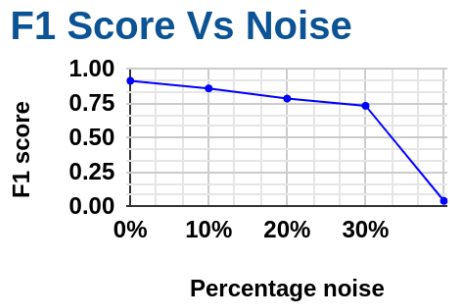
# Experimental results under salt noise insertion attacks

Salt noise is a very common type noise seen in the images. It is also known as impulsive noise. In this part of our experiment, we consider that if an adversary tries to inject the salt noise on the data, it effects the machine learning model significantly.



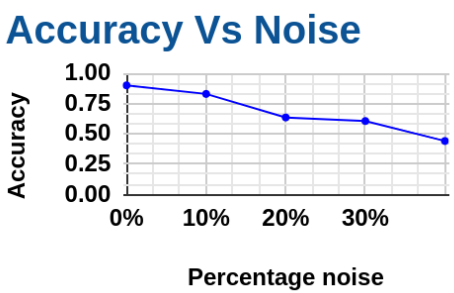



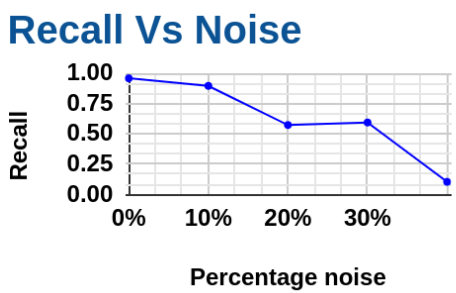



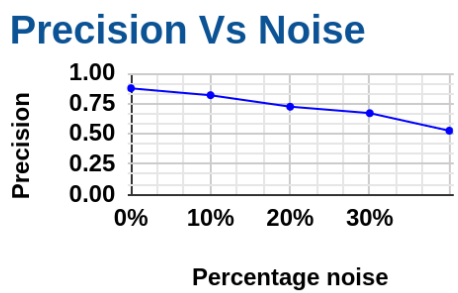



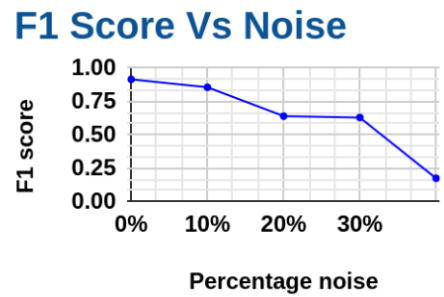
# Experimental results under Gaussian noise insertion attacks

Gaussian noise is another standard noise that has used in our experiments. It is referred as a statistical noise having a probability density function equal to that of the Gaussian distribution.



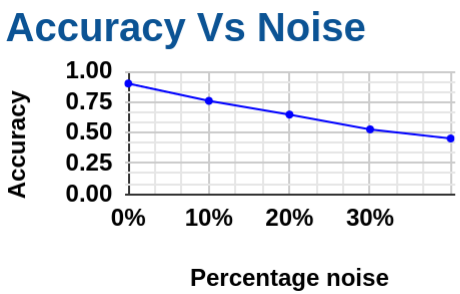



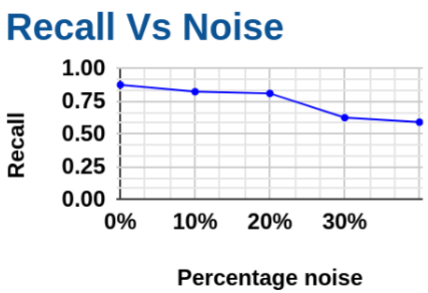



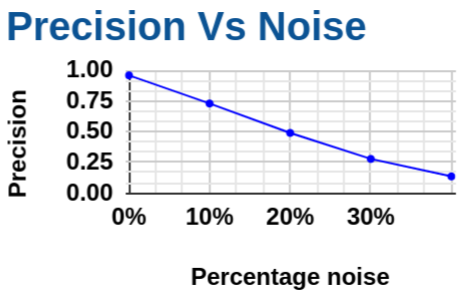



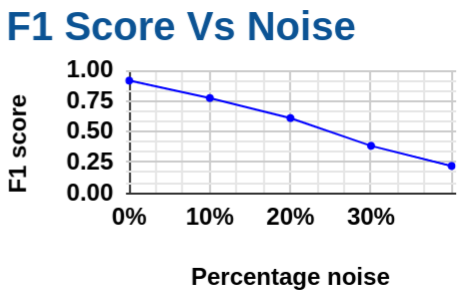
# Experimental results under Poisson noise insertion attacks

Poisson noise is a standard noise that is used in the experimental results to check the effect of accuracy, recall, precision and F1 score under the ML model.



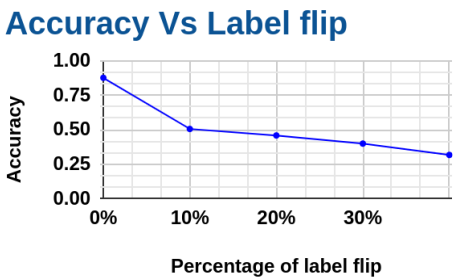



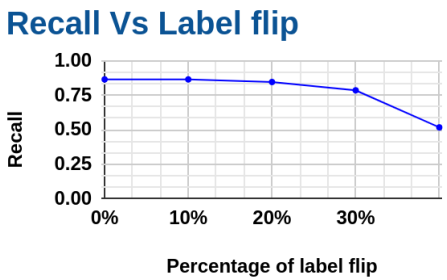



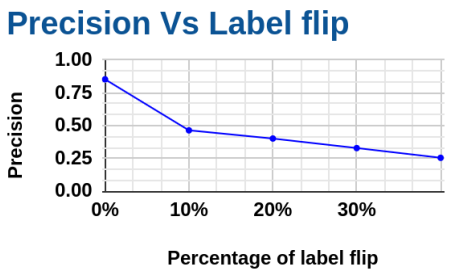



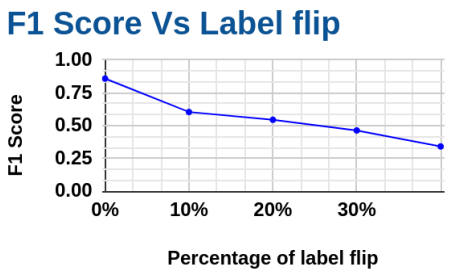
# Experimental results under label flipping attacks

In a label filliping attack, if the adversary tries to alter the labels of the original data, how it can effect on the overall performance of the ML model.



# Performance of ML model without data poisoning attacks



- The effect of no data poising attacks under the ML model.
- 0% noise insertion means that no attacks on the data.

Table: Experimental results without any data poisoning attacks

| Accuracy | Recall | Precision | F1 Score |
|---|---|---|---|
| 0.8764 | 0.8648 | 0.8533 | 0.859 |

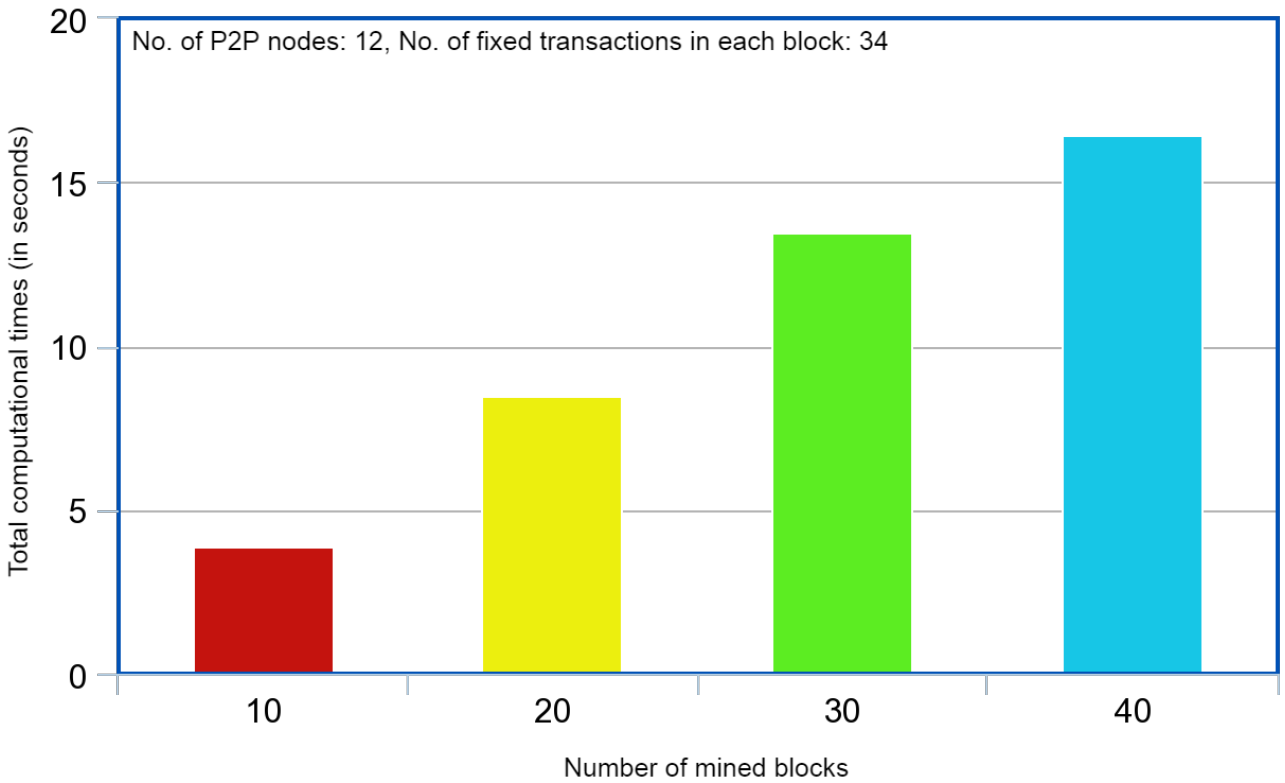
# Blockchain Implementation



- We created a blockchain system with the help of node.js script. We also created the virtual distributed blockchain network with twelve P2P (cloud) servers.
- All the servers work on the localhost, but they use different ports for communication among them.
- To simulate the blockchain system, we have used a host computer having the configuration as: “OS: Ubuntu 18.04 LTS, Processor: Intel i5-8400 (2.80GHz), Memory: 7.6 GiB, OS Type: 64 bit, Disk Type: HDD, Disk Size: 152.6 GB”.
- We have considered two different cases.

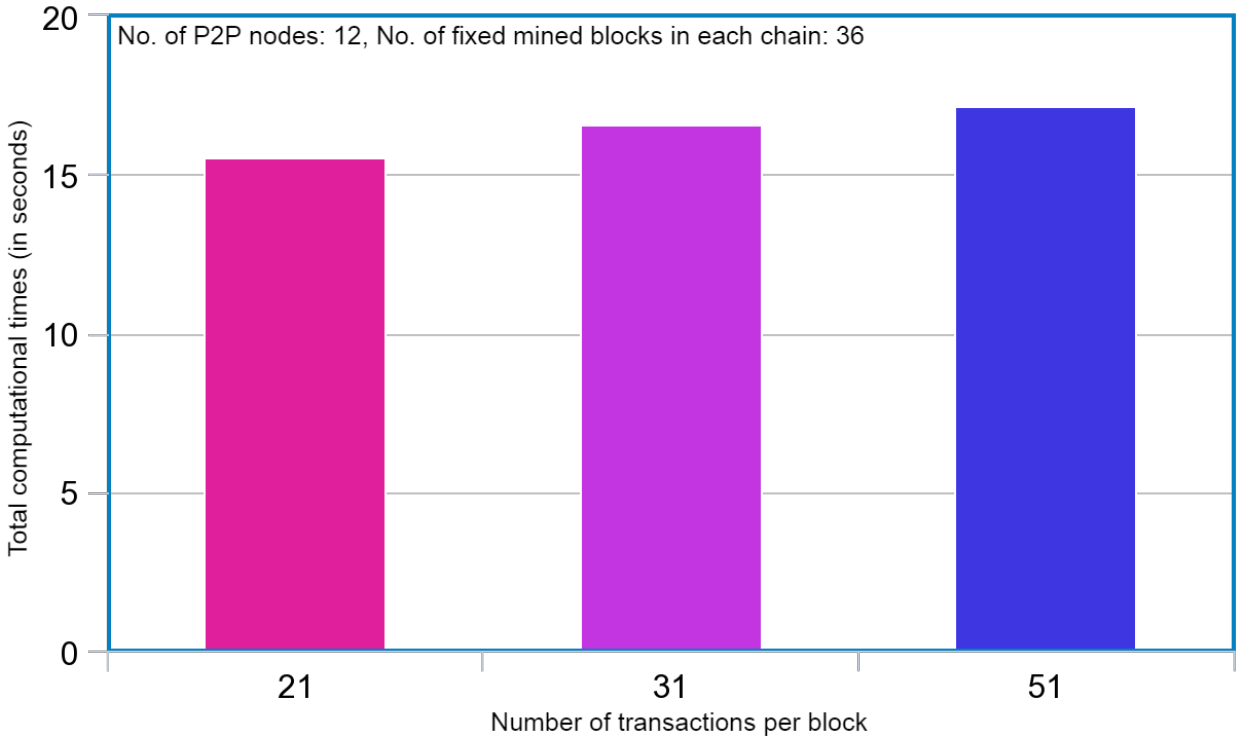
# Blockchain Implementation

**Case 1.** In this case, we fixed the number of transactions to a fixed value as 34 (i.e., the size of each block will be the same) and recorded the time of the blocks addition when we varied the number of blocks in the blockchain. The simulation results show that when the size of each block is fixed, the total computation time linearly varies with the number of blocks mined in the blockchain.



# Blockchain Implementation

**Case 2.** In this case, we fixed the total number of blocks in the blockchain to a fixed value as 36, and recorded the time of the blocks addition when we varied the size of the blocks (i.e., we varied the number of transactions in a block). The simulation results illustrate that when the number of blocks into the blockchain is fixed, the total computation time varies linearly with the number of transactions in each block in the blockchain.



# Thank You
# For Your Attention