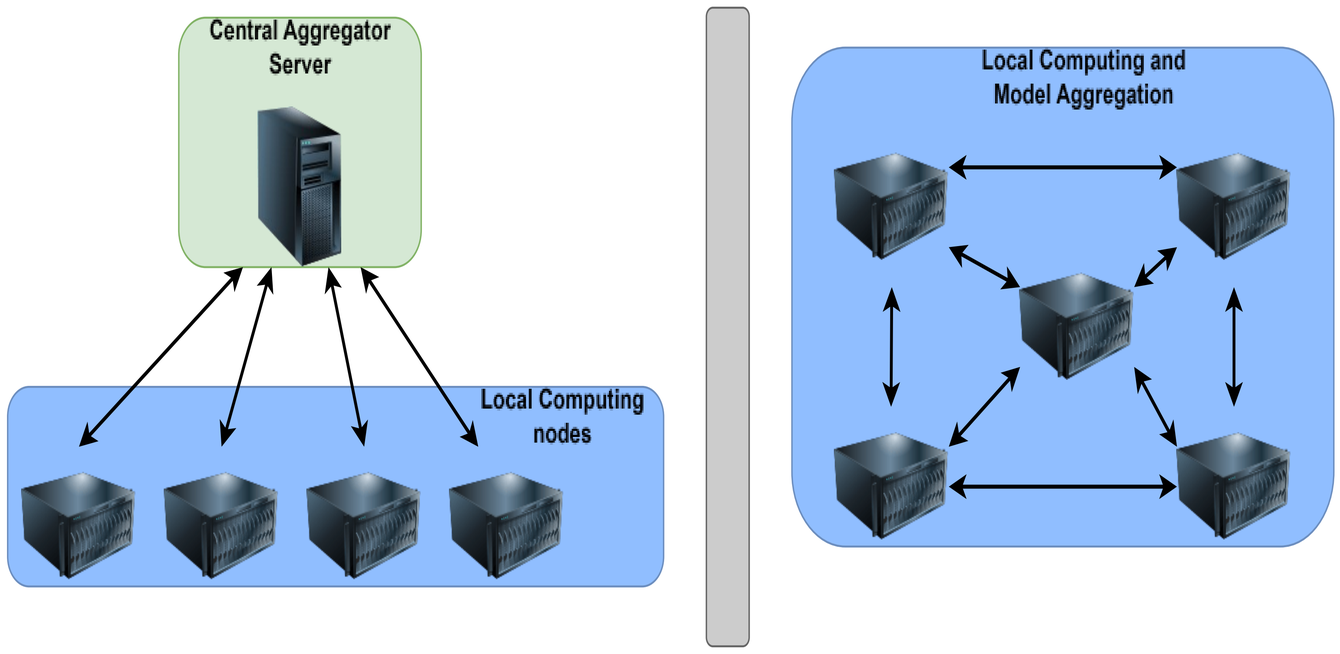
MLOps and distributed machine learning on the edge of the network

By: Ari Kukkaro,
Doctoral student
Email: ari.kukkaro@tuni.fi

Decentralized Machine Learning

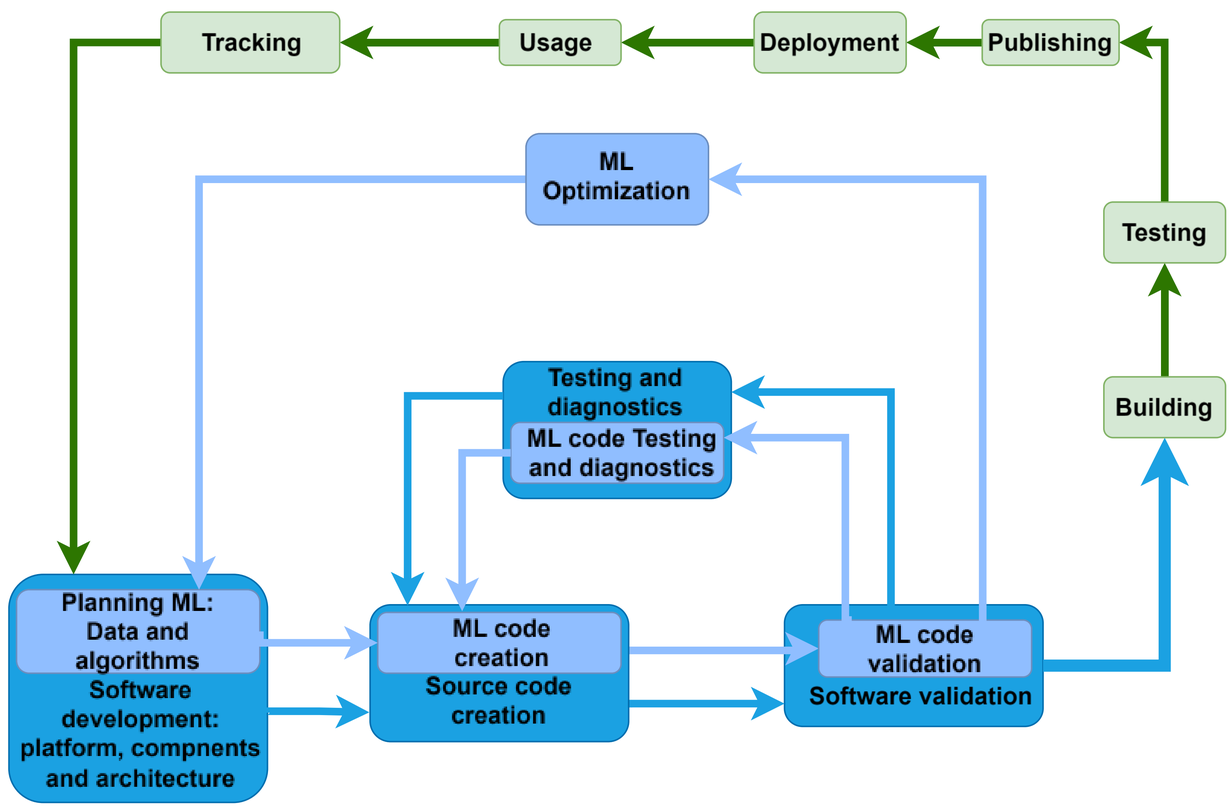
Advances in technology and machine learning have created opportunities to analyse vast amounts of data collected from sensors, industrial devices and systems. Machine learning can be used to assist medical experts in diagnosis, fault detection and accident prevention of industrial equipment. Traditionally, all this data has been collected on centralized servers such as the cloud. Distributed Machine Learning (DML), on the other hand, allows computing tasks to be distributed across multiple computing devices.

Distributing computing tasks and data across multiple devices provides the benefits of large data partitioning, better privacy, security and faster ML model training. Restricting data to a specific device in DML provides privacy and security in design. Then, multiple devices can train a common model with less data per unit than the centralized server would, reducing ML model training time. Federated Learning is a sub-concept of DML in which multiple computing devices collaboratively train a common machine learning model.



MLOps

Machine Learning Operations or MLOps extends the software development framework Development and Operations or DevOps. MLOps provides a flexible, fast and realistic way to develop, test, deploy and maintain ML models as part of the application or software. MLOps provides a defined way to create ML applications and ensure their reliability and updating, which is critical in the development of industrial ML products.

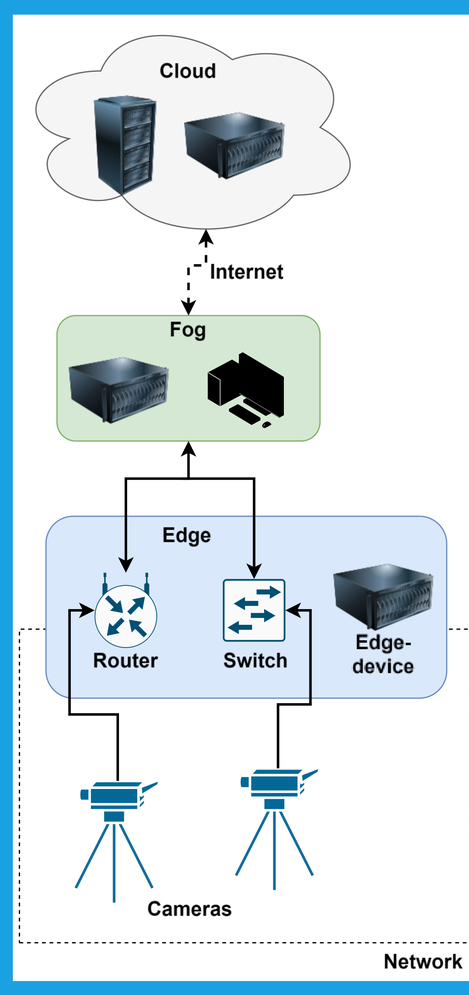


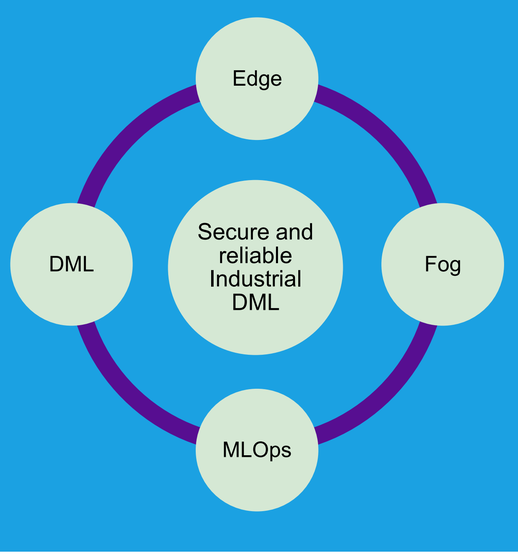
MLOps pipeline adopted from [1]

Research method and data

Multivocal Literature Review (MLR) is first research which aim is to gain knowledge about the Distributed Machine Learning and its sub-concept Federated Learning (FL). Then the rest of the research focuses on the use of constructivist research, which belongs to mixed research methods, combining qualitative and quantitative research methods. The aim of the research is to innovate constructs that answer the research questions on the right.

- **RQ1** How distributed machine learning is applied to industrial settings?
- **RQ2** How to manage securely industrial environment data flows with distributed machine learning?
- **RQ3** How distributed machine learning can be included in efficient and flexible software development processes (DevOps and MLOps)?





Expected Results

The expected results of this work are to gain more knowledge about Distributed Machine Learning and MLOps so that DML can be applied in industrial settings.
Another objective is to investigate how to securely manage DML data flows in an industrial setting. Research will focus on how this can be done using the edge and fog computing paradigms.
The final goal is to bridge the gap between efficient and dynamic industrial software development and secure distributed machine learning.

References

[1] Sergio Moreschini, Franscesco Lomio, David Hästbacka and Davide Taibi, 2022, MLOps for evolvable AI intensive software systems, 2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)