Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 170 (2020) 571–577

Procedia Computer Science

www.elsevier.com/locate/procedia

The 11th International Conference on Ambient Systems, Networks and Technologies (ANT)
April 6-9, 2020, Warsaw, Poland

Open Source Machine Learning Frameworks for Industrial Internet of Things

Asharul Islam Khan [a]*, Ali Al-Badi[b]

[a] *Dept. Operation Management and Business Statistics,*
Sultan Qaboos University, P.O.Box 33 AlKhodh, Muscat, P.C.123, Oman
[b]*Academic Affairs and Research, Gulf College, P.O.Box 885, Al Khuwair, Al Mabaila, Muscat, P.C.133*

Abstract

Information and communication technology has revolutionized the industrial operations and productions. The industries irrespective of size, whether small or large, have felt the need of artificial intelligence and machine learning techniques to process the terabytes of data generated through sensors, actuators, industrial management systems, and web applications. These data have the characteristics of volume (terabyte) and variety (image, audio, video, graphics) and thus customized models and techniques are required for analysis and management.

The advancement in computer hardware, processing power, storage capacity, and cloud computing have led to experimentation and implementation of machine learning models in industrial domain for resource optimization, operation management, and quality control. However, the industrial Data Analysts face the dilemma of selecting the affordable and easy to use machine learning frameworks that suite their need and expectations. The study investigates the open source machine learning frameworks, aligned with the industrial domain (processing data generated from Industrial Internet of Things), in terms of usage, programming languages, implementations, and future prospects.

Keywords: industrial internet of things; artificial intelligence; machine learning; deep learning; industrial 4. revolution; open source; frameworks; sensors; Tensorflow.

* Corresponding author. Tel.: +968-95687668.
E-mail address: ashar.367@gmail.com

10.1016/j.procs.2020.03.127

1. Introduction

Machine learning applications are quickly transforming the industrial landscape. Many businesses have reduced the production and operation costs using tools powered by machine learning models and algorithms. The deep learning which is a subset of machine learning has found ways in manufacturing, industrial maintenance, drug discovery, pattern imaging analytics, and software testing [1]. The deep learning a type of deep neural network consisting of layered structure as input layer, hidden layer, and output layer.

Industrial Internet of Things (IIoT) is defined as a set of machines, robotics, cognitive technologies, and computers for intelligent industrial operations with the help of data analytics [2]. The Industrial Internet of Things is a part of Industry 4.0 revolution, which is concerned with automation, innovation, big data, and cyber physical systems in industries. The Industrial Internet of Things are showing positive impact in supply chain, transportation, healthcare, manufacturing, oil and gas, energy/utilities, chemical, and aviation industry. The Industrial Internet of Things has helped in controlling and monitoring manufacturing and production from remote locations [3-5]. The Industrial Internet of Things market will reach $123.89 Billion by 2021 [6].

Industrial Internet of Things captures large chunk of data, later used for predictive maintenance, time management, and cost control after machine learning models implementation. The machine learning models forms the core of logistics and supply chain solutions in terms of optimizing the product packet size, delivery vehicle selection, delivery route selection, delivery time computation. For instance DHL uses Amazon's Kiva robotics (improve speed, accuracy) for the network management.

The Industrial Internet of Things and machine learning models are inseparable entities for optimal solutions as far as the industrial context is concerned. However, the machine learning models need development, training, and testing in a software/ programming framework before being put in actual use. These software/ programming framework (IBM Watson) are often termed as software development environment/ model development environment, have licensed fee. The licensed fee prevents small industries in experimenting the machine learning models for their own need. Therefore, the study illustrates the open source machine learning frameworks (Tensorflow, Torch, etc.) for designing machine learning model using data generated through Industrial Internet of Things. Even small size industries can experiment with machine learning models for business forecasting and resource management.

There are five sections. Section 2 corresponds to research background and motivation. Section 3 illustrates the open source machine learning frameworks. Section 4 presents results and discussion. The last section concludes the findings and future research.

2. Research Background and Motivation

According to one estimate, the number of Industrial Internet of Things devices will increase to 75.44 billion by 2025. The technologies enabling Industrial Internet of Things are: Internet of Things, artificial intelligence, cloud computing, artificial intelligence for Cyber Physical Systems (CPS), big data analytics, blockchain, augmented and virtual reality[7]. The key elements of Industrial Internet of Things are the smart devices, communication network, and big data analytics, Figure 1 shows the key elements of Industrial Internet of Things.

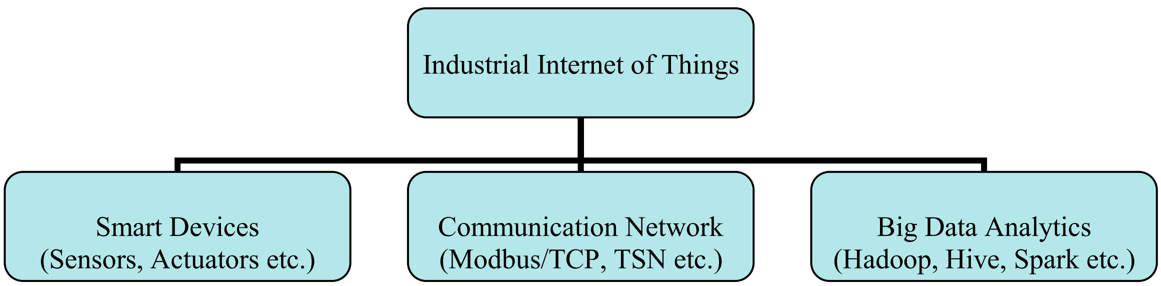

Fig 1. Key Elements of Industrial Internet of Things

The machine learning is a subset of artificial intelligence. The machine learning is of various types such as supervised (input and output features are known in advance), unsupervised learning (algorithm/ model learns dataset patterns and group them into clusters), and reinforcement learning (learning takes place because of reward and punishments). The image classification, fraud detection, disease diagnosis, weather, and market forecasting uses supervised learning (Linear Regression, Support Vector Machine, Decision Tree, K Nearest Neighbor). The recommender system and big data visualization uses unsupervised learning (K-Mean Clustering, FP Growth). The real time decision and robot navigation uses reinforcement learning.

The machine learning technologies in the form of mobile intelligences and Automated Guided Vehicles (AGVs), are used within industries to improve productivity, efficiency and reliability [8, 9]. The Industrial Internet of Things applications in industries requires standards and intelligence techniques such as low-power wireless networking technologies and fast sensors for big data analytics [10]. The data collected during industrial production and operation by the sensors are stored over cloud for analysis and prediction [11]. The machine learning algorithms are used in the feature selection of Industrial Internet of Things [12].

The machine learning models are trained and tested on the data sets generated by the Industrial Internet of Things for smart decision [13]. Machine learning has become an essential part of Industrial Internet of Things for efficient and effective industrial outputs. Industrial units face uncertain problems of equipment failure and downtime. The machine learning implementation possibly prevents and improves the unexpected system outcomes. The trained and tested machine learning models in the industrial settings contribute to predictive maintenance and risks reduction. The energy sector needs to know the power consumption and future demands based on the historical data, population parameters, and weather conditions. The machine learning models optimizes product market fit, identifies anomalies in manufacturing, manages operational risk, prevents churn, and enhances communications.

Nevertheless, to implement machine learning models one need to select the machine learning frameworks. The following section provides the open source machine learning frameworks and implementation process in an industrial setting.

3. Open Source Machine Learning Frameworks and Implementation Process

The design, development, training, and testing of a machine learning model depends on the choice of machine learning frameworks. There are numbers of machine learning frameworks for developing the machine learning models. The open source frameworks for the machine learning model development in an industrial environment is shown in Figure 2.

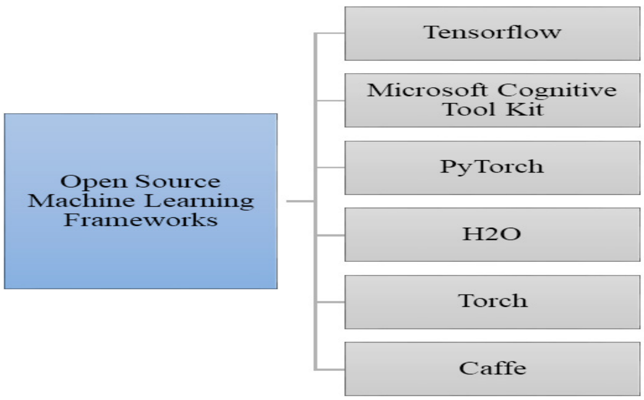

Fig 2. Open Source Machine Learning Frameworks for Industrial Internet of Things

Tensorflow: The Google owns Tensorflow. It is used in deep learning (deep learning is a subset of machine learning). The Tensorflow has many versions but the latest versions are faster, more flexible, and support new languages. It is used by multinational companies such as Qualcomm (for snapdragon mobile platforms), Intel (for

Intel platform), China mobile (for detecting network anomalies'), and CEVA (deep learning processors). Tensorflow supports high performance computation for face and handwriting recognition. It run on Graphical Processing Unit (GPU), CPU, servers, desktops, and mobile devices. The C++/ Python graphs from Tensorflow is used for processing on CPUs or GPUs.

Microsoft Cognitive Toolkit: The Microsoft owns the Microsoft Cognitive Toolkit. It helps businesses and organizations in exploring the machine learning solutions. It is an open source deep learning development environment and supports multi-machine-multi-GPU-back-ends. Initially, it was designed to learn on the pattern of human brain, now it can be used for feed-forward Deep Neural Network (DNN), Recurrent Neural Networks (RNN), and Convolution Neural Networks (CNN). Its usage include predictive maintenance solutions of aircrafts. In September 2017, Microsoft released Azure Machine Learning Model Management tool to help developers in management and deployment of machine learning workflows.

H2O: It is the one of the biggest open source machine learning model development framework for business and enterprises. The HORTIFRUT industry optimizes the processing of blue berries using H2O. The Stanley Black and Decker optimizes the manufacturing processes and reduces the time and costs. The Intel uses it for network traffic and intrusion detection.

Torch: The Torch is an open source machine learning model development framework. The Facebook owns it. The Torch is a library, scientific computing framework, and script language for machine learning based on GPU/CPU. Torch is flexible and easy to use. It is used by companies such as Purdue, Yandex, and NVIDIA.

PyTorch: The Facebook owns PyTorch. It is similar to Torch and less matured than Tensorflow. The PyTorch follows an object-oriented paradigm. The code writing in PyTorch is easy due to conditionals and loops functionality. The PyTorch is used by companies such as IBM, Facebook, and Yandex.

Caffe (Convolutional Architecture for Fast Feature Embedding): It was developed at Berkeley Vision and Learning Center (BVLC). It is one of the fastest systems for DNN and can process 60M images/day with a single GPU. Caffe offers easy configuration and switching between GPU and CPU for model training. The companies such as Google and Pinterest uses Caffe for computer vision, speech, and multimedia content analysis. The Caffe, is used to build Convolutional Neural Network (CNN) for image classification.

Machine learning based systems in manufacturing sector solves the supply chain and predictive maintenance problems. The Airbus has a dedicated unit for identifying gaps in the manufacturing process by using sensors, tools, and wearable technologies. Figure 3 shows the machine learning frameworks implementation process in industrial setting.

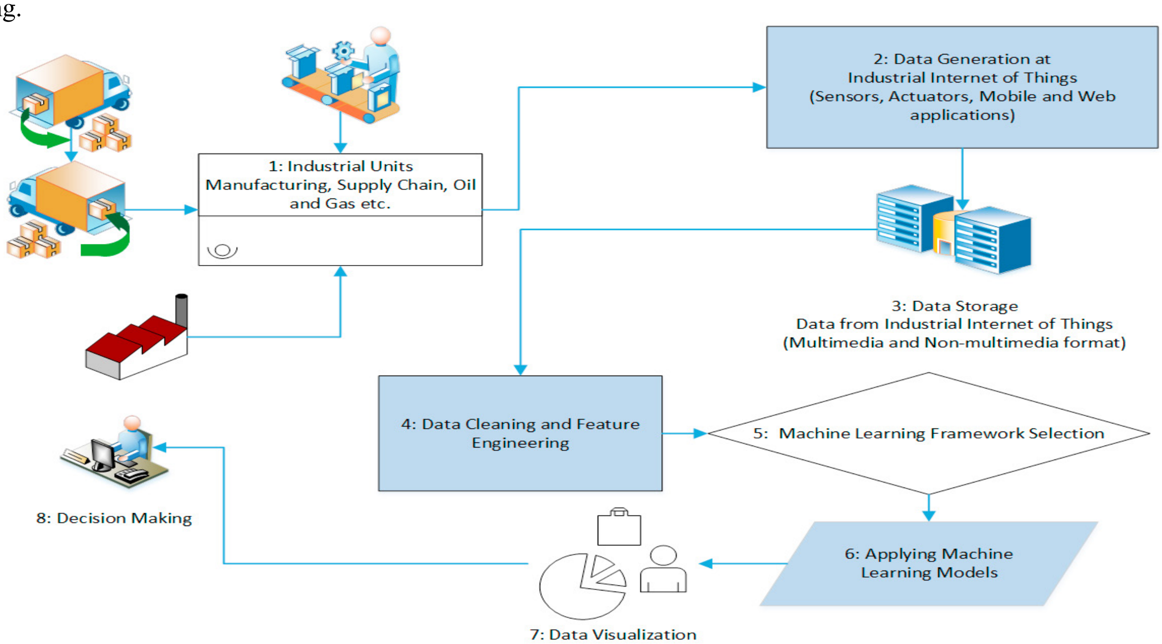

Fig 3. Machine Learning Framework Implementation Process in an Industrial Setting

The industries such as manufacturing, oil and gas, and supply chain generates data through sensors, actuators, and mobile applications. These data are stored either at local server/ at cloud for further processing. After cleaning the data, the machine learning frameworks such as Tensorflow, Torch etc., selected for applying different machine learning models (Linear Regression, Support Vector Machine, Decision Tree, etc.) on the given datasets. The model producing the highest score in the confusion matrix for a given data set is accepted and it is used for prediction on a new data set. The decision makers uses the machine learning model outputs in decision-making process.

4. Results and Discussion

The machine learning models provide solutions to industry specific issues in the form of revenue generation, customer behavior investigation, production process automation, and price optimization besides prediction and forecasting. The machine learning models help in anomaly detection (steel manufacturing plants), quality assessment (automobile industry), and inventory management in in manufacturing industries.

The Industrial Internet of Things (sensors, actuators, web applications, and management systems) generates huge volume and variety of data. The volume and variety are big data characteristics [14]. In the first instance, the data is used for developing, training and testing of a machine learning model. The machine learning algorithms (classification (Random Forest, Support Vector Machines), clustering (K-Means, K-Nearest Neighbors), and binary classification (Logistic Regression)) which fits best with the data forms the base of the machine learning model. The machine learning frameworks help in the development of machine learning model. Table 1 shows the open source machine learning frameworks.

Table 1. Comparison of Open Source Machine Learning Model Development Environment

Machine Learning Frameworks	Supporting Language	Supporting Platform	Maintainer	Applicability
Tensorflow	Python and C++	Windows, MacOS and Linux	Google	Deep speech, smart reply, and computer vision.
Microsoft Cognitive Toolkit	Python, C++, and C#	Windows and Linux	Microsoft Research	Handwriting, image, and voice recognitions
Caffe	C, C++, Python, and MATLAB	Ubuntu, MacOS, and Windows	BVLC	Training models for classification.
H2O	Java and Python	Windows and Linux	H2O	Creates productionize machine learning models.
Torch	Lua	Linux, Android, MacOS, iOS and Windows	Ronan, Clément, Koray and Soumith	Detecting and solving hardware problems for data flows.
PyTorch	Python	Linux, Android, MacOS, iOS, and Windows.	Ronan, Clément, Koray and Soumith	Reinforcement learning and scaled production of models.

Although, the manufacturing and transportation companies are spending millions of dollars on Industrial Internet of Things and its services, not every industry with Industrial Internet of Things has been able to exploit the generated data and machine learning applications. It is important to choose the machine learning frameworks considering interoperability with Industrial Internet of Things data. Figure 4 shows the machine learning frameworks and the trends on Google during last one year.

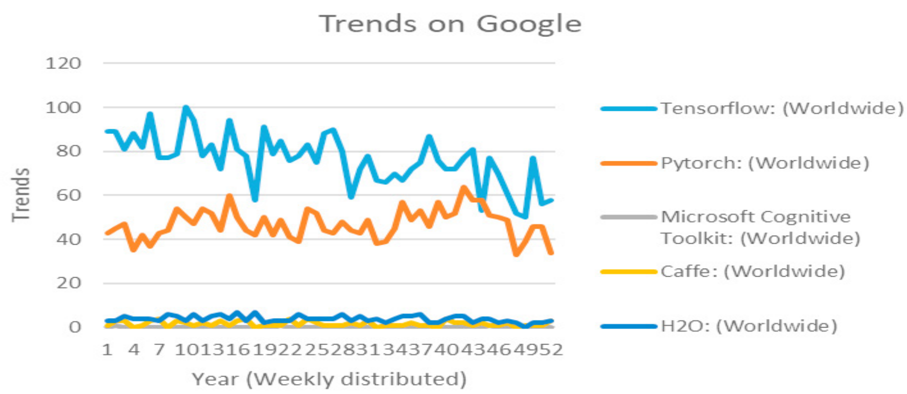

Fig 4. Machine Learning Frameworks and Google Trends during last year (2019)

Among the available open source machine learning frameworks, the Tensorflow has the largest community and user base. The PyTorch has the second largest community. The Microsoft Cognitive Toolkit, Caffe, and H2O have very low visibility among the users. The Tensorflow and PyTorch are on the top among the five framework, as they support all the operating systems with documentation. Among the other things these two frameworks support speech recognition and computer vision (object detection, object recognition, object classification, object segmentation). The speech recognition and computer vision is the hottest topic of discussion and research in industrial environments.

5. Conclusion

Machine learning are shaping the future industries. The machine learning has the potential to innovate the industries and their business models. The machine learning implementation on Industrial Internet of Things data has tremendous benefits to industries like consumer demand prediction in energy industries, supply chain adjustments, predictive maintenance, quality control, and increase production throughput.

However, new entrants face difficulties to understand and use the available machine learning frameworks for data processing and machine learning model development. It is important to investigate the available open source machine learning frameworks for reaping the benefits. The study has presented five open source frameworks for machine learning model development. The study further shows the data generation in Industrial Internet of Things environment and the machine learning framework implementation process.

The future work would implement Tensorflow on Industrial Internet of Things data from oil and gas industry to develop a predictive maintenance model based on machine learning and deep learning.

References

[1] Khan, Asharul Islam and Yaseen Al-Mulla. *Unmanned Aerial Vehicle in the Machine Learning Environment*. in *10th International Conference on Emerging Ubiquitous Systems and Pervasive Networks*. 2019. Portugal: Procedia Computer Science.

[2] Sadiku, Matthew NO, Yonghui Wang, Suxia Cui, and Sarhan M Musa, *Industrial internet of things.* International Journal of Advances in Scientific Research and Engineering, 2017. **3**.

[3] Patel, Alpesh, Rohit Singh, Jignesh Patel, and Harsh Kapadia. *Industrial Internet of Thing Based Smart Process Control Laboratory: A Case Study on Level Control System*. in *International Conference on Information and Communication Technology for Intelligent Systems*. 2017: Springer.

[4] Mumtaz, Shahid, Ai Bo, Anwer Al-Dulaimi, and Kim-Fung Tsang, *Guest editorial 5G and beyond mobile technologies and applications for industrial IoT (IIoT).* IEEE Transactions on Industrial Informatics, 2018. **14**(6): p. 2588-2591.

[5] Yao, Haipeng, Tianle Mai, Jingjing Wang, Zhe Ji, Chunxiao Jiang, and Yi Qian, *Resource Trading in Blockchain-based Industrial Internet of Things.* IEEE Transactions on Industrial Informatics, 2019.
[6] iscoop. *The Industrial Internet of Things (IIoT): the business guide to Industrial IoT*. 2018 [cited 2019 10 june]; Available from: https://www.i-scoop.eu/internet-of-things-guide/industrial-internet-things-iiot-saving-costs-innovation/.
[7] Khan, WZ, MH Rehman, HM Zangoti, MK Afzal, N Armi, and K Salah, *Industrial internet of things: Recent advances, enabling technologies and open challenges.* Computers & Electrical Engineering, 2020. **81**: p. 106522.
[8] Duan, Ying, Yun Luo, Wenfeng Li, Pasquale Pace, Gianluca Aloi, and Giancarlo Fortino, *A collaborative task-oriented scheduling driven routing approach for industrial IoT based on mobile devices.* Ad Hoc Networks, 2018. **81**: p. 86-99.
[9] Hassanzadeh, Amin, Shimon Modi, and Shaan Mulchandani. *Towards effective security control assignment in the Industrial Internet of Things*. in *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*. 2015: IEEE.
[10] Dujovne, Diego, Thomas Watteyne, Xavier Vilajosana, and Pascal Thubert, *6TiSCH: deterministic IP-enabled industrial internet (of things).* IEEE Communications Magazine, 2014. **52**(12): p. 36-41.
[11] Ma, Mimi, Debiao He, Neeraj Kumar, Kim-Kwang Raymond Choo, and Jianhua Chen, *Certificateless searchable public key encryption scheme for industrial internet of things.* IEEE Transactions on Industrial Informatics, 2017. **14**(2): p. 759-767.
[12] Zhao, Long and Xiangjun Dong, *An industrial Internet of Things feature selection method based on potential entropy evaluation criteria.* IEEE Access, 2018. **6**: p. 4608-4617.
[13] Li, YangQun, *An Integrated Platform for the Internet of Things Based on an Open Source Ecosystem.* Future Internet, 2018. **10**(11): p. 105.
[14] Al-Badi, Ali, Ali Tarhini, and Asharul Islam Khan, *Exploring Big Data Governance Frameworks.* Procedia Computer Science, 2018. **141**: p. 271-277.