Using AI on IoT Sensor Data - for predicting health of man and machine

Dr. Arpan Pal
arpan.pal@tcs.com
Chief Scientist and Research Area Head
Embedded Systems and Robotics, TCS Research
Tata Consultancy Services, India

Introduction

The Internet of Things (IoT) is pervading everyday life of both end-consumers and enterprises as the main technology for “Digitalization” [1]. With the cost of sensors coming down, cloud technology becoming more mature, capacity of internet becoming better and computing platforms on both cloud and edge becoming faster / cheaper / more available, the whole system is poised for a disruptive impact across industries. Different Industry verticals are seeing different business value of using IoT using the context-aware “Sense-Analyze-Respond” cycle [2], where Sensing is via IoT, Analyze is via Artificial Intelligence (AI) and Respond is either via human-in-loop visualization or via robotic actuation.

In an earlier article in IEEE India Council Newsletter (Apr-Jun 2018) [3], we had presented “Cognitive IoT”, which discussed about promised benefits of marriage of IoT with AI. As that promise becomes reality in practice, the true business benefits of such systems are emerging along with some practical problems that need solving. In this article we present real-life use cases for AI-based analytics on IoT data and try to elaborate the practical problems associated with such implementations. We focus on two industry verticals – Healthcare and Manufacturing as use case exemplars.

Motivating Use Cases

According to World Health Organization (WHO), non-communicable diseases (NCD) are becoming the biggest silent killer of humanity – of every 100 deaths occurring in the world (including accidents), 71 of them are caused by NCD like cardiovascular disorders, diabetes, old age problems, neurological disorders and musculoskeletal disorders [4]. The main problem with NCD is the fact that they are very often asymptomatic in their early stages – by the time symptoms are manifested, the disease has already progressed. On the other hand, if the disease was detected in early stage, simple lifestyle changes are often enough to control the disease. Most of these diseases have very well-established gold-standard specific tests to diagnose the disease onset (ECG analysis and coronary angiogram for cardiovascular problems [5] like atrial fibrillation and coronary artery disease, psychometric tests for detection of dementia and mild cognitive impairments [6] in elderly people). But the problem is that people don’t know they have the disease in the early asymptomatic stage and hence they don’t consult a doctor for having these tests done as a screening mechanism. Hence there is a need to create AI-based screening algorithms using easily available data from wearable sensors and home medical devices, which can be comfortably used at home as a screening mechanism for the above-mentioned conditions that can suggest whether to consult a doctor.

In the manufacturing world, there is a similar problem of predicting machine health that can help in more efficient predictive maintenance of machines [7]. A specific problem in this space is prediction of the “remaining-useful-life” (RUL) [8] of machines and machine parts using various sensor data like vibration, current load, heat / sound generated etc. AI driven predictive analytics of the sensor data followed by multi-sensor fusion can yield reasonably high accuracy for RUL prediction. Similarly, real-time control of process parameters to improve the quality of the final product based on sensing the product quality is also an important use case for manufacturing [9]. Creation of such control algorithms can immensely benefit from AI driven inferencing based on learning from past data.

Implementation Examples

There are now very good instances of detection of atrial fibrillation (AF) from single lead ECG data. There was a global challenge arranged by Physionet [10] in which people used supervised deep learning and traditional machine learning based algorithms to achieve nearly 85% accuracy in detecting AF [11], [12]. Similarly significant accuracy has been reported for detection of coronary artery disease using supervised machine learning based classifiers (85% using only heart sound from digital stethoscope [13], 95%+ by augmenting the heart sound with electrocardiogram (ECG), photoplethysmogram (PPG) from pulse oximeter and patient family history knowledge [14].

Similarly on RUL estimation of machines, mean average percentage error (MAPE) of 19% has been reported with False Positive Rate (FPR)in real-life datasets using supervised machine learning techniques [15] [16]. An interesting example of real-time process control using AI and IoT is the case of Friction-stir-welding quality improvements using Friction-stir Welding machines where physics based modelling followed by machine learning based techniques can yield improved results [17].

Practical Points to Ponder

As AI based analytics of IoT data starts proliferating the industry, there are some interesting practical issues cropping up that need attention. Many of these are open questions which the practitioners need to ponder over. We list a few of them here.

1) *Trust and Liability of AI-based Inferencing* – It is very unlikely to create AI-based inferencing that is 100% accurate. But this brings in an inherent unpredictability in the behavior of the machine deploying the AI. This raises a bigger question – as human beings are we ready to trust machines that are intrinsically unpredictable [18]? Who will take the liability of decisions made by AI inferencing? For example, if AI predicts a person to have certain kind of disease, who will take the liability of false positive and false negatives? In all such scenarios, AI should not be seen as an independent inferencing system, but an aid to human-in-loop decision making – in the healthcare example, we can think of such an AI system helping a doctor to take a more informed decision about the patient with liability of the decision remaining with the doctor.
2) *Interpretable AI* - The idea of human-in-loop inferencing using AI brings in another issue – will the human experts (like doctors) be comfortable with AI systems that don't have interpretable models? Without that, the experts cannot relate the AI driven inferencing to available body of scientific knowledge (like medical knowledge / machine design knowledge). Many of today's deep learning systems on sensor data cannot provide such interpretability and hence may have acceptance issue in the expert community as a human-in-loop inferencing aid [19]. As has been highlighted in the recent ICML workshop on Human Interpretability in Machine Learning [20], *"Supervised machine learning models boast of remarkable predictive capabilities. But can you trust your model? Will it work in deployment? What else can it tell you about the world? We want models to be not only good, but interpretable. And yet, the task of interpretation appears underspecified."*
3) *Understanding the Sensor Signal Morphology* – Deep learning based AI systems has shown very good results on human-generated data like images, text or speech (which has an inherent structure governed by some rules like language etc.). But the same cannot be said yet for sensor signal analytics systems where lacks a defined structural model due to varied signal morphology. As pointed out by Prof. Josh Tanenbaum, at MIT [21], *"There's no way you can have an AI system that's human-like that doesn't have language at the heart of it. It's one of the most obvious things that set human intelligence apart."* When we deal with raw sensor data generated naturally by a machine or human body, such structure is lacking, thereby making it difficult for deep learning systems to derive value from it. Hence there is need to augment deep learning based algorithms with traditional signal processing based approach for sensor time series data – such hybrid approaches are already yielding good results [22].
4) *Non-availability of data and labels* – Availability of sufficient data to train AI models is always a problem which is more pronounced in deep learning based systems. Even if data is collected and models are trained for one scenario (may be one type of factory or health data from people a particular country), there is no guarantee that such a trained model will work in a different but similar scenario (another factory with similar machines or another country with different demography people). This raises some few very important but practical aspects –
    a. Few-shot learning [23] and Meta-learning [24] – Systems should be able to learn quickly on a few instances of training data and should be able to use meta-knowledge available to augment the data learning.
    b. Unsupervised learning and Transfer learning – Systems should be able to infer reasonably in the absence of labels or where labelling can be done on demand by human experts on a reduced subset of the data identified by unsupervised approach. Transfer learning techniques can help in re-training existing pre-trained models from one scenario dataset with a small representative data from the new scenario.
5) *AI at the edge* – Edge devices / on-premise devices play a large role in IoT systems. In the context of AI based analytics, they play significant role to provide
    a. Low-latency, real-time inferencing needed for IoT-driven process control systems.
    b. Low-battery consumption that is needed for energy-constrained devices like wearable and implantable.
    c. Privacy-preserving analytics as the data does not leave the edge / premises even for analytics.

    AI at edge either needs special technique to compress the AI models enabling them to run on constrained edge devices, or have dedicated low-latency, low memory inferencing algorithms, or have special purpose hardware accelerators in the edge [25], [26]. In order to reduce the energy consumption by significant order, completely new processor architectures called neuromorphic [27], that mimic the brain in hardware, are been used to design new chips - such chips, coupled with new sensing techniques called spiking sensors and 3rd generation brain-inspired neural networks called spiking neural networks (SNN) [28] hold the promise of disrupting the low-power edge AI technology.

Conclusion

Creating value via AI based analytics of IoT sensor data has started showing promise in real world deployments. However, there are quite a few practical challenges outlined in this article that needs to be addressed before it creates disruptive impact. In this article, we have tried to look into few of those challenges with specific use case examples in Healthcare (AF and CAD detection) and Manufacturing (RUL estimation and real-time process control). These challenges include Trust/Liability, Interpretability, Signal Morphology understanding, non-availability of data/labels and Edge Computing.

References

[1] Parida, Vinitn et. al., "Reviewing literature on digitalization, business model innovation, and sustainable industry: Past achievements and future promises.", MDPI Sustainability 2019, 11(2), 391
[2] Arpan Pal et. al.., “IoT Technical Challenges and Solutions”, Artech House, Dec. 2016
[3] https://site.ieee.org/indiacouncil/files/2018/08/icnl-18q2.pdf
[4] https://vizhub.healthdata.org/gbd-compare/
[5] https://www.medicinenet.com/coronary_angiogram/article.htm
[6] https://emedicine.medscape.com/article/1136393-overview
[7] https://towardsdatascience.com/predictive-and-prescriptive-maintenance-for-manufacturing-industry-with-machine-learning-2078afa76bfb
[8] https://medium.com/@RemiStudios/remaining-useful-life-in-predictive-maintenance-ffc91d7e4a97
[9] https://www.isa.org/intech/20181202/
[10] https://physionet.org/content/challenge-2017/1.0.0/
[11] Gary Clifford et. al., “AF Classification from a short single lead ECG recording: thePhysioNet/Computing in Cardiology Challenge 2017”, Computing in Cardiology (CinC), 2017
[12] Mukherjee, Ayan et. al. "Detection of atrial fibrillation and other abnormal rhythms from ECG using a multi-layer classifier architecture." Journal of Physiological measurement 40, no. 5 (2019)
[13] https://physionet.org/content/challenge-2016/1.0.0/
[14] Banerjee, Rohan et. al. "Non-Invasive Detection of Coronary Artery Disease Based on Clinical Information and Cardiovascular Signals: A Two-Stage Classification Approach." IEEE International Symposium on Computer-Based Medical Systems (CBMS), 2018
[15] Gugulothu, Narendhar et. al. "Predicting remaining useful life using time series embeddings based on recurrent neural networks." arXiv preprint arXiv:1709.01073 (2017).
[16] C. Zhang et. al., "Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations", Sensors, vol. 16, no. 6, 2016.
[17] Prashant Chauhan et. al. “Modeling of defects in friction stir welding using coupled Eulerian and Lagrangian method”, Journal of Manufacturing Processes, Elsevier Publication, May, 2018.
[18] https://sloanreview.mit.edu/article/can-we-solve-ais-trust-problem/
[19] https://www.tcs.com/blogs/role-of-artificial-intelligence-in-predictive-analysis-using-sensor-data
[20] https://arxiv.org/pdf/1606.03490.pdf
[21] https://www.technologyreview.com/s/602094/ais-language-problem/
[22] Arijit Ukil et. al., “Fusing Features based on Signal Properties and TimeNet for Time Series Classification”, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), 2019
[23] https://medium.com/quick-code/understanding-few-shot-learning-in-machine-learning-bede251a0f67
[24] https://medium.com/huggingface/from-zero-to-research-an-introduction-to-meta-learning-8e16e677f78a
[25] https://www.wired.com/story/microsoft-sends-a-new-kind-of-ai-processor-into-the-cloud/
[26] https://www.intel.ai/intel-neural-compute-stick-2-smarter-faster-plug-and-play-ai-at-the-edge/
[27] https://www.technologyreview.com/s/526506/neuromorphic-chips/
[28] https://www.frontiersin.org/articles/10.3389/fncom.2018.00042/full

About the author

Arpan Pal received both his B.Tech and M.Tech from Indian Institute of Technology, Kharagpur, India in Electronics and Telecommunications and PhD. from Aalborg University Denmark. He is a Senior Member of IEEE and is engaged in the innovation space in different industry bodies and start-up accelerators.

He has more than 26 years of experience in the area of Signal Processing, Communication, Embedded Systems and Robotics. Currently he is with Tata Consultancy Services (TCS), where, as Chief Scientist, he is heading the Embedded Systems and Robotics Research Area in TCS Research. His research interests include Unobtrusive Sensing, AI-driven Sensor Signal Informatics & Edge Processing for applications in Connected Health, Smart Cities, Smart Enterprises & Industry 4.0; and also Cognitive Robotics & Collaborative Drones for Telepresence & Disaster Handling.

Arpan has more than 140 publications and book chapters till date in reputed Journals and Conferences. He has also authored a complete book on IoT. He has filed for more than 100 patents and has 60 patents granted to him. He has been on the editorial board for reputed journals like ACM Transactions on Embedded Computing Systems, IEEE Transactions on Emerging Topics in Computing and IT Professional Magazine from IEEE Computer Society.

Home Page: http://research-innovation.tcs.com/research/Pages/Pal-Arpan.aspx
LinkedIn: http://in.linkedin.com/in/arpanpal
Google Scholar: https://scholar.google.co.in/citations?user=hkKS-xsAAAAJ&hl=en