A CRITICAL REVIEW ON AI ENABLED IOT SERVICES

Amit Kumar Sharma, Assistant Professor, Department of Computer Applications, Galgotias University

ABSTRACT

In a coming years, billions of linked gadgets will be deployed around our homes, towns, cars and industries in the globe. Devices with limited resources interact with the environment and individuals around them. Many of these gadgets use machine learning models to decipher significance and behaviour underlying the data of sensors, to execute precise predictions and make judgments. The bottleneck is the high amount of linked elements that may congestion the network. Therefore, intelligence on end devices must be incorporated with machine learning algorithms. The deployment of machine learning on such edge devices reduces the network congestion by enabling computation near the data sources. The objective of this work is to evaluate key strategies guaranteeing the execution of machine-learning models on hardware with low performance in the paradigm of the Internet of Things, paving the path to the Internet of Conscious Things. In this work we give a complete assessment of models, architecture and requirements for solutions implementing cutting edge machine learning on the internet of Things devices, with the primary purpose of defining state-of-the-art and imagining development needs. An example of the implementation of edge machine learning on a microcontroller, widely known as "Hello World" machine learning, is presented.

Keywords: Artificial Intelligence, IOT, Machine Learning, Human

INTRODUCTION

The Internet of Things (IoT) scenario has been well recognised in recent years. It has a software and hardware architecture which connects the physical world to the Internet. In recent years, the significant rise in interest in this paradigm has dramatically increased the number of IoT devices. By 2025, nearly 75 billion devices are expected to be connected to the internet, with an economic impact on the global economy. IoT devices normally have limited calculating power and little memory, which may create huge data volumes. Low power and interconnected systems, notably sensors, are used in our homes, cities, cars and enterprises. While cloud computing might be suitable for IoT growth, along with the projected bandwidth saturation the data transmission latency is not sufficient for some applications (e.g. wellness monitoring). With the growing number of connected devices, cloud processing alone might become

impractical and result in higher latency, decreased bandwidth and secrecy and dependability. The calculation must thus be performed locally as possible by exploiting intelligence on terminals to minimise cloud traffic. This means giving gadgets a form of "consciousness" which may interact in the absence of a connection, thereby producing sophisticated behaviours and the "Internet of Conscious Things" adaptability to quickly changing settings. Unfortunately, processing limitations on resource-scarce devices hamper advanced machine learning (ML) approaches, although several software-based frameworks offer reliable and cost-efficient ways of optimising separate edge computing installations. Tasks which may be supplied for edge elements include low data fusion while efficient systems must be allocated to a more in-depth understanding of the data (e.g. decision-making purposes). However, the transfer of raw data to cloud servers increases connection costs, produces a delayed system response, and exposes private data. One practical option is to consider local processing of data to their origins and transport just the data necessary for further cloud processing to remote servers. Edge computing indicates that computations are performed as close to data sources rather than distant places as possible.

MACHINE LEARNING

As shown in earlier reports, ML algorithms may be implemented on devices with low calculating power, thereby allowing edge computing to be employed to better the IoT area. A similar phrase fog computing proposes an architecture in which the cloud is stretched towards the IoT-end devices, enhancing latency and security by executing computations close to the edge of the network. Therefore, even for fog calculation, the aim is to bring the processing phase where the data is created closer, but where the "intelligence" is placed is the primary difference. In fog computing the processing stage is in a fog node or in an IoT gateway at level LAN (local area network). Data are usually processed immediately on the sensor devices in edge computing (physically very close to the sensors). The closer to the sensor it is the better it is in terms of privacy and energy consumption since the energy need for the transmission of data is reduced. In such situation, a range of energy harvesting options from various sources opens the way for passive or semipassive IoT sensor systems AI-enabled.

DEEP LEARNING

A profound learning model is a mixture of weights and prejudices. Those parameters vary according to an optimization function that assesses the prediction capacity of the model based on the objective function (loss function or reward function if the study is respectively supervised or strengthened). After a training stage, the AI system determines the underlying pattern between the data and predicts a value based on the input data. There are various options for designing the Neural Network model if various network hyperparameters provide a varying degree of accuracy. In instance, a high-precision model demands more memory than a low-accuracy model owing to the amount of parameters. The metric utilised for precision measurement relies on the scope of the application of the ML algorithm. In object detection, for example,

precision may be quantified by an average mAP which indicates how well the predicted item's position aligns with the ground-truth position, averaged across several item categories.

MODEL COMPRESSION

Model compression enables us to execute the model on small devices and there are two major approaches to minimise the network: lower accuracy (low bits per weight) and less weight (pruning). Post-training quantization minimises demand on computer power and energy to the detriment of a low accuracy loss. The model weights are float32 variables by default, which cause two problems: Firstly, the model is rather huge since 4 bytes have a significant memory demand in each weight; secondly, the performance is shockingly sluggish compared to the Uint8 type variables. Weights may be significantly reduced from 32 to 8 bits, therefore reducing the size of the NN 4 times. Note that post-quantization is a method that is done after the model is trained, but may be done before the training. ML packages like Tensor flow or Keras allow the use of quantization. As indicated above, model size reduction may be achieved not only by quantizing, but also by cutting procedures to remove connections that cannot help the NN; this results to a reduction in the calculation demand and programme memory. Quantization and trimming procedures both separately and collectively were investigated.

AI AND WIRELESS EMBEDDED IOT SYSTEM

The communication protocol waste most of the energy in the IoT device. In fact, any unneeded information that is sent, stored and processed seems to constitute potential energy waste. Excellent algorithms must thus also be supported by effective protocols of communication. Developers might employ several communication protocols according to the unique circumstances. This is because we can characterise protocols that enable us to transfer a little quantity of data with a low energy consumption over long distances and protocols that are capable of transmitting large amounts of data over long distances with a high consumption. If the usage of spectrum bands has been taken into account, we may also categorise them in technologies which utilise the licenced or unlicensed spectrum, e.g. the ISM bands. Among the various Communication Technologies, BLE provides an example of how a smart sensor with 2,4 GHz Radio Frequency (RF) power and an infrared motion and BLE communication is designed and optimised. Wireless Bluetooth technology is extensively utilised, including introducing Bluetooth 5 which consumes less power and allows mesh topology, provides broad networks for appliances and multiple communications. With its strong range, faster speed up to 2Mbps and a long-range, improved sensitivity with fewer bits, Bluetooth 5 fits the standards for new IoT devices.

CONCLUSION

By enabling calculations to be done nearby to data sources, ensuring privacy while uploading data, and lowering power consumption for continuous wireless transmission to gateways or cloud servers, deploying machine learning on Internet of Things devices alleviates network congestion. The purpose of this study

was to explore the primary strategies for executing machine learning models on low-performance hardware in the Internet of Things paradigm, paving the way for the Internet of Conscious Things. The purpose of this study was to provide a comprehensive evaluation of the models, architectures, and needs for solutions that utilise edge machine learning on IoT devices, with the primary objective of defining the state of the art and conceptualising development needs.

REFERENCES

1. Ahmadi, A. R., Shahbazi, S., & Diyanat, M. (2016). Efficacy of Five Herbicides for Weed Control in Rain-Fed Lentil (Lens culinaris Medik.). *Weed Technology*, *30*(2), 448–455. https://doi.org/10.1614/WT-D-15-00125.1

2. Gagic, V., Hänke, S., Thies, C., Scherber, C., Tomanović, Ž., & Tscharntke, T. (2012). Agricultural intensification and cereal aphid-parasitoid-hyperparasitoid food webs: Network complexity, temporal variability and parasitism rates. *Oecologia*, *170*(4), 1099–1109. https://doi.org/10.1007/s00442-012-2366-0

3. Michta, E., Szulim, R., Sojka-Piotrowska, A., & Piotrowski, K. (2017). IoT-based flood embankments monitoring system. In R. R. S. Linczuk M. (Ed.), *Proceedings of SPIE - The International Society for Optical Engineering* (Vol. 10445). SPIE. https://doi.org/10.1117/12.2280830

4. Petracca, M., Bocchino, S., Azzarà, A., Pelliccia, R., Ghibaudi, M., & Pagano, P. (2013). WSN and RFID integration in the IoT scenario: An advanced safety system for industrial plants. *Journal of Communications Software and Systems*, *9*(1), 104–113. https://doi.org/10.24138/jcomss.v9i1.162

5. 85062942572&partnerID=40&md5=82aca5b6254ea6c5d214d08f61aa85d3

6. Valente, F. J., & Neto, A. C. (2017). Intelligent steel inventory tracking with IoT / RFID. *2017 IEEE International Conference on RFID Technology and Application, RFID-TA 2017*, 158–163. https://doi.org/10.1109/RFID-TA.2017.8098639

7. Chiuchisan, I., Costin, H.-N., & Geman, O. (2014). Adopting the internet of things technologies in health care systems. In H. C.-G. Gavrilas M. Ivanov O. (Ed.), *EPE 2014 - Proceedings of the 2014 International Conference and Exposition on Electrical and Power Engineering* (pp. 532–535). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICEPE.2014.6969965

8. Chung, P.-T., Chung, S. H., & Hui, C.-K. (2012). A web server design using search engine optimization techniques for web intelligence for small organizations. *2012 IEEE Long Island Systems, Applications and Technology Conference, LISAT 2012*. https://doi.org/10.1109/LISAT.2012.6223208

9. Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., Blondel, V. D., & Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the*

National Academy of Sciences of the United States of America, *111*(45), 15888–15893. https://doi.org/10.1073/pnas.1408439111

10. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056
11. Forbes, E. E., Hariri, A. R., Martin, S. L., Silk, J. S., Moyles, D. L., Fisher, P. M., Brown, S. M., Ryan, N. D., Birmaher, B., Axelson, D. A., & Dahl, R. E. (2009). Altered striatal activation predicting real-world positive affect in adolescent major depressive disorder. *American Journal of Psychiatry*, *166*(1), 64–73. https://doi.org/10.1176/appi.ajp.2008.07081336
12. Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. *International Journal of Medical Informatics*, *101*, 75–84. https://doi.org/10.1016/j.ijmedinf.2017.02.002
13. Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Martonosi, M., Rowland, J., & Varshavsky, A. (2011). Identifying important places in people's lives from cellular network data. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6696 LNCS*, 133–151. https://doi.org/10.1007/978-3-642-21726-5_9
14. Jenaro, C., Flores, N., Gómez-Vela, M., González-Gil, F., & Caballo, C. (2007). Problematic internet and cell-phone use: Psychological, behavioral, and health correlates. *Addiction Research and Theory*, *15*(3), 309–320. https://doi.org/10.1080/16066350701350247
15. Junco, R., & Cotten, S. R. (2012). No A 4 U: The relationship between multitasking and academic performance. *Computers and Education*, *59*(2), 505–514. https://doi.org/10.1016/j.compedu.2011.12.023
16. Klausen, J. (2015). Tweeting the Jihad: Social media networks of Western foreign fighters in Syria and Iraq. *Studies in Conflict and Terrorism*, *38*(1), 1–22. https://doi.org/10.1080/1057610X.2014.974948
17. Lester, R. T., Ritvo, P., Mills, E. J., Kariri, A., Karanja, S., Chung, M. H., Jack, W., Habyarimana, J., Sadatsafavi, M., Najafzadeh, M., Marra, C. A., Estambale, B., Ngugi, E., Ball, T. B., Thabane, L., Gelmon, L. J., Kimani, J., Ackers, M., & Plummer, F. A. (2010). Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): A randomised trial. *The Lancet*, *376*(9755), 1838–1845. https://doi.org/10.1016/S0140-6736(10)61997-6
18. Liang, Y., Reyes, M. L., & Lee, J. D. (2007). Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, *8*(2), 340–350. https://doi.org/10.1109/TITS.2007.895298
19. Lowres, N., Neubeck, L., Salkeld, G., Krass, I., McLachlan, A. J., Redfern, J., Bennett, A. A., Briffa, T., Bauman, A., Martinez, C., Wallenhorst, C., Lau, J. K., Brieger, D. B., Sy, R. W., & Ben Freedman, S. (2014). Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies: The SEARCH-AF study. *Thrombosis and Haemostasis*, *111*(6), 1167–1176. https://doi.org/10.1160/TH14-03-0231

20. Mateo, G. F., Granado-Font, E., Ferré-Grau, C., & Montaña-Carreras, X. (2015). Mobile phone apps to promote weight loss and increase physical activity: A systematic review and meta-analysis. *Journal of Medical Internet Research*, *17*(11). https://doi.org/10.2196/jmir.4836
21. Morone, F., & Makse, H. A. (2015). Influence maximization in complex networks through optimal percolation. *Nature*, *524*(7563), 65–68. https://doi.org/10.1038/nature14604
22. Morris, M. E., Kathawala, Q., Leen, T. K., Gorenstein, E. E., Guilak, F., Labhard, M., & Deleeuw, W. (2010). Mobile therapy: Case study evaluations of a cell phone application for emotional self-awareness. *Journal of Medical Internet Research*, *12*(2). https://doi.org/10.2196/jmir.1371
23. Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., & Zhou, C. (2014). A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science*, *28*(9), 1988–2007. https://doi.org/10.1080/13658816.2014.913794
24. Pierannunzi, C., Hu, S. S., & Balluz, L. (2013). A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004-2011. *BMC Medical Research Methodology*, *13*(1). https://doi.org/10.1186/1471-2288-13-49
25. Poggio, T., & Bizzi, E. (2004). Generalization in vision and motor control. *Nature*, *431*(7010), 768–774. https://doi.org/10.1038/nature03014
26. Potlapally, N. R., Ravi, S., Raghunathan, A., & Jha, N. K. (2006). A study of the energy consumption characteristics of cryptographic algorithms and security protocols. *IEEE Transactions on Mobile Computing*, *5*(2), 128–143. https://doi.org/10.1109/TMC.2006.16
27. Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design*, *33*(5), 727–748. https://doi.org/10.1068/b32047
28. Ahn, Y.-Y., Bagrow, J. P., & Lehmann, S. (2010). Link communities reveal multiscale complexity in networks. *Nature*, *466*(7307), 761–764. https://doi.org/10.1038/nature09182
29. Aranda-Jan, C. B., Mohutsiwa-Dibe, N., & Loukanova, S. (2014). Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC Public Health*, *14*(1). https://doi.org/10.1186/1471-2458-14-188
30. Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis and Prevention*, *71*, 311–318. https://doi.org/10.1016/j.aap.2014.06.005
31. Case, M. A., Burwick, H. A., Volpp, K. G., & Patel, M. S. (2015). Accuracy of smartphone applications and wearable devices for tracking physical activity data. *JAMA - Journal of the American Medical Association*, *313*(6), 625–626. https://doi.org/10.1001/jama.2014.17841
32. Cerf, M., Paxon Frady, E., & Koch, C. (2009). Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal of Vision*, *9*(12), 1–15. https://doi.org/10.1167/9.12.1
33. Chen, G.-H., Chen, W.-Y., Yen, Y.-C., Wang, C.-W., Chang, H.-T., & Chen, C.-F. (2014). Detection of mercury(II) ions using colorimetric gold nanoparticles on paper-based analytical

devices. *Analytical Chemistry*, *86*(14), 6843–6849. https://doi.org/10.1021/ac5008688

34. Khan N, Yaqoob, Hashem, et al. Infrared Spectroscopy of Metal-MgO Single Crystalline Composite Films. *Jpn J Appl Phys*. 2014;29(8):2-17. doi:10.1143/JJAP.29.L1497
35. Frumkin H, Hess J, Luber G, Malilay J, McGeehin M. Climate change: The public health response. *Am J Public Health*. 2008;98(3):435-445. doi:10.2105/AJPH.2007.119362
36. Lee D, Kang S, Shin J. Using deep learning techniques to forecast environmental consumption level. *Sustain*. 2017;9(10):1-17. doi:10.3390/su9101894
37. Parry ML, Canziani O., Alcaraz E de A, et al. *Climate Change 2007: Impacts, Adaptation and Vulnerability*.; 2007. doi:10.1016/B978-008044910-4.00250-9
38. Kumar GS, Kar SS, Jain A. Health and environmental sanitation in India: Issues for prioritizing control strategies. *Indian J Occup Environ Med*. 2011;15(3):93-96. doi:10.4103/0019-5278.93196
39. Levac D, Colquhoun H, O'Brien KK. Scoping Studies: Advancing the Methodology. *Implement Sci*. 2010;5(1):2-9. doi:10.1017/cbo9780511814563.003
40. Summers JK, Smith LM, Fulford RS, Crespo R de J. The Role of Ecosystem Services in Community Well-Being. *Ecosyst Serv Glob Ecol*. 2018. doi:10.5772/intechopen.74068