Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 71–78
DOI: 10.15439/2022F267
ISSN 2300-5963 ACSIS, Vol. 31

Distributed and Adaptive Edge-based AI Models for Sensor Networks (DAISeN)

Veselka Boeva, Emiliano Casalicchio
Shahrooz Abghari
Ahmed A. Al-Saedi, Vishnu Manasa Devagiri
Computer Science Department, Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden
Email: {vbx, emc}@bth.se

Andrej Petef
Peter Exner, Anders Isberg, Mirza Jasarevic
Sony Europe BV
R&D Center Europe
Lund, Sweden
Email: andrej.petef@sony.com

***Abstract*—This position paper describes the aims and preliminary results of the *Distributed and Adaptive Edge-based AI Models for Sensor Networks* (DAISeN)[1] project. The project ambition is to address today's edge AI challenges by developing advanced AI techniques that model knowledge from the sensor network and the environment to support the deployment of sustainable AI applications. We present one of the use cases being considered in DAISeN and review the state-of-the-art in three research domains related to the use case presented and directly falling into the project scope. We additionally outline the main challenges identified in each domain. The developed Global Navigation Satellite Systems (GNSS) activation model addressing the use case challenges is also briefly introduced. The future research studies planned for the remaining period of the project are finally outlined.**

I. Introduction

THE NUMBER of solutions that provide Artificial Intelligence (AI) and Machine Learning (ML) based systems has been growing recently. These solutions facilitate the creation of new smart products and services in many different fields. In addition, sensor networks are undergoing great expansion and development and the integration of AI and sensor networks benefits many areas such as Industry 4.0, healthcare, mobility, logistics, and many other Internet-of-Things (IoT) applications. However, this has also put new challenges in front of researchers and practitioners. New real-time AI and ML algorithms are needed along with different strategies to embed these algorithms in sensor boards and network nodes such as fog/edge nodes. For example, edge-based AI requires robust and adaptive models that take into account the temporal component of a data flow and allow for vertical and horizontal scaling of the decision-making process. These models must employ efficient learning algorithms that are capable of dealing with information varying over time and coping with large scale missing and inaccurate values. In addition, the decision-making models should be composable so that they can be distributed on the edge devices in order to ensure a trade-off between the decision accuracy, latency, and consumed energy per decision.

The IoT is an emerging key technology for future industries and the everyday lives of people, e.g., it has been playing an increasingly important role in healthcare, agriculture, home services, industrial processes, and transportation. Wireless Sensor Network (WSN) is an enabling technology for IoT [1], and, by definition, is the bridge between the physical world and the intelligence residing on the Internet. The integration of AI and sensor networks (by means of the IoT) are now realities that are changing our lives. Sensor networks are widely used to collect environmental parameters, e.g., in homes, buildings, and vehicles, where they are used as a source of information that aids the decision-making process and, in particular, it allows systems to learn and monitor activity. Although there are numerous advantages of sensor networks, it should be mentioned that they also consume energy and contribute to E-waste [2]. These place new stress on the environment and the smart world.

According to the World Economic Forum (WEF)[2], the IoT is undoubtedly one of the largest enablers for responsible digital transformation. WEF survey has outlined that more than 80% of IoT deployments are currently addressing, or have the potential to address the Sustainable Development Goals (SDGs)[3] defined by the United Nations, for example, industry, innovation, and infrastructure; smart cities and communities; affordable and clean energy (SDG #7); good health and well-being (SDG #3); clear water and sanitation (SDG #6); smart agriculture; responsible production and consumption (SDG #12).

Today's IoT solutions embed and leverage AI both in the end-user services and the network's management. To boost sustainability, IoT solutions need to be sustainable and usable. These goals are achievable only by means of advances in AI, decision making, and edge and fog computing. AI algorithms and decision-making models are at the core of state-of-the-art and future IoT applications and need to be distributed among IoT devices, such as WSN nodes, and edge and fog

This work is part of the Sony RAP 2020 Project "Distributed and Adaptive Edge-based AI Models for Sensor Networks"

[1]Daisen is a volcanic mountain located in Tottori Prefecture, Sanin Region of Japan.

[2]World Economic Forum, Internet-of-Things Guidelines for Sustainability (2018) http://www3.weforum.org/docs/IoTGuidelinesforSustainability.pdf

[3]United Nations' Sustainable Development Goals https://sdgs.un.org/goals

nodes, to scale vertically and horizontally, and to minimize energy consumption. At the same time, the management of complex IoT infrastructure (that includes fog WSN and edge/fog computing nodes) should be operated with energy-aware decision-making mechanisms that leverage distributed ML/AI techniques. The DAISeN project aims to address the challenges discussed above by developing:

- advanced AI algorithms for continual, shared, and evolving learning, that enable learning from multiple data sources by distributed training, and continual updating of the model;
- distributed decision-making models allowing vertical and horizontal scaling in order to guarantee high-quality decision-making concerning time consumption, energy consumption, and data communication.

This position paper is organized as in what follows. The research objectives of the DAISEeN project are described in Section II. In Section III-B we outline one of the use cases being considered in DAISeN and review the state-of-the-art ML and data mining techniques applicable at the edge. They are analyzed in relation to the use case to bring to evidence the challenges and gaps the project aims to fill. Preliminary results are reported in Section IV. Finally, the paper is concluded by our outlook (Section V).

II. Research Objectives

The main focus of the proposed research lies in the usability of AI on edge devices and fog nodes to improve the performance and sustainability of sensor networks and the training process.

The DAISEeN project will investigate methods for transferring and adapting AI algorithms to edge devices with limited computational performance. The ambition is the development of resource-aware AI algorithms that can be run on edge nodes. These algorithms take into account the hardware and software capabilities of the edge nodes and the capabilities of the communication links between these nodes, always keeping in mind that a balance is found between the limited energy resources of edge devices and the complexity of the AI algorithms. Therefore, the network traffic between the edge devices has to be kept low. That can be achieved by several methods such as context-aware techniques [3], dynamic device clustering, role assignments, or intelligent sensor fusion, and data reasoning techniques that can account for dynamically changing surrounding environment, including context prediction.

Another challenge addressed by DIASeN is the development of advanced AI algorithms for continual, shared, and evolving learning that enable learning from multiple data sources by distributed training and continual updating of the model. This can be achieved by developing unsupervised and semi-supervised methods to automate knowledge extraction and learning in data stream scenarios [4], [5]. The main problem investigated is how the newly arrived information can be taken into account in the learning phase and can be used for continuous adaptation of the learned model [6]. In addition, it will be studied how to develop, train, and evaluate a model with no direct access to labeled data. Candidate approaches to address those challenges are:

1) dynamic unsupervised and semi-supervised learning models that are robust to the appearance of drifting context and additionally enable to learn from multiple data sources by distributed training, and continual updating and evolving of the model [4], [6], [7], [8];
2) development of dynamic techniques for automatic annotation (labeling) of the data;
3) usage of transfer learning techniques enabling reuse of knowledge from training in earlier tasks to subsequent tasks.

The other research ambition of DAISeN is the design of distributed/composable data mining models. These models will allow to vertically/horizontally scale the decision making in order to guarantee high quality decisions in edge computing environments. This can be achieved by adopting ML models that can predict the computational level (i.e., cloud vs. edge) with respect to the network operational context (e.g., latency vs. accuracy). Such models will allow one to run the lighter but less accurate models at the edge for the sake of latency, and the computation-intensive but higher-accuracy models in the cloud. The focus is on developing data mining and ML techniques to maintain local models embedded in the edge devices and further integrate low-level edge devices' observations into a global model [4], [5], [9]. Computed at a higher level, the latest can produce reliable decisions based on the available input data. Those techniques will produce a global model even when data from some devices are missing due to network changes or degradation.

III. DAISeN Use Case and State of the Art

A. *A context-aware GNSS activation use case*

Sony provides software solutions in smart logistics for monitoring and tracking goods. GNSS is the used positioning technique for detecting the tracker's current position. GNSS is known to perform well in open sky environments. However, the trackers may be in any place, such as open outdoors, crowded city areas, indoors, and so on. Sony would like to perform context-aware control of GNSS activation by automatically and accurately detecting indoor/outdoor localization of the trackers by consuming the least energy. The Sony requirement is to use radio signals received from Long-Term Evolution (LTE) base stations to detect the environment (indoor/outdoor). The main idea is that the propagation of radio signals is affected by the environment. Different environmental scenarios have different signal strength characteristics. By learning different signal strength characteristics, it would be possible to determine the tracker's environment.

The setup explained above requires advanced AI solutions that can detect indoor/outdoor localization (environmental context) of the tracker for controlling GNSS activation in order to save power. In addition, these models are expected to be able to continuously adapt to new scenarios and environments, as well as to learn in the distributed framework from many trackers to improve environmental context detection.

B. Current state of the art and challenges

In this section, we review ML and data mining techniques related to the use case described in Section III-A. The discussed techniques fall into three main research domains: context-awareness at the edge, continual and evolving learning, and federated learning, see Table I.

1) Context awareness at the edge: Our considerations are limited to indoor/outdoor context-awareness methods at the edge due to the fact that it is related to the use case discussed above (see Section III-A).

Efficient knowledge discovery is critical for the optimized operation and management of IoT/sensor networks. Context may affect the complex systems' operation and management procedures at various levels, from the physical to the communication, up to the application level [10]. For example, as discussed above, positioning edge devices, such as smartphones and trackers, in outdoor areas typically rely on GNSS such as the Global Positioning System (GPS), which performs well in open sky environments. However, these devices may be in any place, such as deep indoors, metal containers, crowded urban areas, etc. In addition, the GPS consumes too much energy to be useful for many applications. Therefore, detecting indoor/outdoor and providing this context-aware information in various environments may be helpful and lead to battery-saving solutions. Many indoor/outdoor detection methods have already been proposed.

In [11], these methods are classified into two main groups: threshold-based techniques and ML-based techniques. Approaches in the first group use fixed detection rules and thresholds, such as a sensor reading above a certain value, to classify an edge node state (e.g., indoor/outdoor). The second group of solutions uses ML algorithms to detect indoor/outdoor status based on features extracted from smartphones, edge nodes, and in general, embedded sensors. ML-based indoor/outdoor environment detection techniques are in the focus of our interest and are reviewed further in this section.

In [12], an approach, entitled SenseMe, uses the C4.5 algorithm on data generated from GPS, gyroscope, accelerometer, and the Bluetooth module to sense environmental context and the context-aware location. In [13], the authors propose a sound-based indoor/outdoor detection method that utilizes binary classification of the environment's acoustic reverberation features. Canovas et al. [14] employ a binary classification technique on the Received Signal Strength Indicator (RSSI) from 802.11 access points to identify a pedestrian's indoor or outdoor status. Ashraf et al. [15] propose MagIO, a solution that utilizes magnetic field signals sensed by smartphones for detecting indoor/outdoor states. Magnetic field features are classified with different ML algorithms, including Support Vector Machines (SVM), Gradient Boosting Machines (GBM), Random Forest (RF), k-Nearest Neighbor (kNN), and Decision Trees (DT). In [16], the authors apply an ML algorithm to classify the neighboring GSM station's signal in different environments and identify the users' current context by signal recognition. Radu et al. [17] propose to detect indoor/outdoor context by employing co-training according to the feature of light, magnetic, and cell sensors. The proposed solution can automatically learn characteristics of new environments and devices, thereby providing a detection accuracy exceeding 90% even in unfamiliar circumstances. Multiple contextual features are also used in [18], which leveraged J48 and other ML algorithms to detect the indoor/outdoor state with high accuracy. An interesting hybrid solution that integrates unsupervised and supervised algorithms relying on the location accuracy and signal strength is introduced in [19].

Challenges: Most of the reviewed approaches rely on the presence of a large amount of labeled data and report higher performance on datasets coming from the same locations/devices as those used to build the model than on new environments. However, when detecting environmental (indoor/outdoor) context at the edge level in real-world scenarios usually labeled data is scarce or entirely missing. Furthermore, the environment dynamic and the context complexity should be taken into account, but at the same time, keeping in mind that the detection models should be light in order to be able to run on the device [3]. Evidently, novel context-aware data mining and learning techniques are needed. These must be resource-efficient, but also should be able to support continual learning from multiple sources and robust model adaptation to new environments.

2) Continual and evolving learning: The main ideas depicted in the continual learning paradigm are knowledge sharing, adaptation, and transfer [20]. Continual learning algorithms may have to deal with catastrophic forgetting [21], data distribution shifts [22], or imbalanced or scarce data problems [23]. Catastrophic forgetting [21] refers to a model experiencing performance degradation at previously learned concepts when trained sequentially in learning new concepts. The catastrophic forgetting is a significant challenge to tackle in the continual learning context since, by definition, the continual learning setting deals with sequences of classes or tasks. Other challenges that should be considered are data distribution shifts and the emergence of new classes. Changes in the data distribution over time are commonly referred to as concept drift. Gepperth and Hammer [22] define two kinds of concept drift: virtual and real. Virtual concept drift concerns the input distribution and may due to imbalanced classes over time. On the contrary, real concept drift is caused by novelty on data or new classes and can be detected by its effect on, e.g., classification accuracy. The continual learning model has to detect the change and automatically fix it. An undetected shift in the data distribution will lead to forgetting. Online change detection algorithms deal with this challenge as it is shown in [24], [25].

The study [20] surveys different supervised continual learning approaches and classifies them into three main categories: replay, regularization-based, and parameter isolation methods. This classification is based on how task-specific information is stored and used throughout the sequential learning process.

TABLE I
2.2. DISTRIBUTED AI REQUIREMENTS IN RELEVANT STATE-OF-THE-ART AREAS

Domain	State-of-the-art area	Relevant use case requirements
Computations at the edge	Context awarness at the edge	Detect indoor/outdoor localization of the edge node (tracker)
Learning from streaming data	Continual and evolving learning	Continuously update the model when new data are available
Distributed AI	Federated learning	Distributed learning from many edge nodes (trackers) for improving the context detection

The most important studies published in these three categories are summarized in Table II. Replay methods replay previous task samples while learning a new task to alleviate forgetting. The replayed samples are either reused as inputs for rehearsal or for constrained optimization of the new task loss to prevent previous task interference. Rehearsal methods explicitly retrain a limited subset of stored samples while training on new tasks [26], [27], [28], [29]. In the absence of previous samples, pseudo rehearsal is an alternative strategy used in early works with shallow neural networks [30], [31], [32], [33]. Constrained optimization is considered an alternative solution to rehearsal by leaving more freedom for backward/-forward transfer. Rehearsal might be prone to overfitting the subset of stored samples and appears to be bounded by joint training [34], [35]. Regularization-based methods introduce an additional term in the loss function, consolidating previous knowledge when learning new data. These methods can further be divided into data-focused and prior-focused methods. Knowledge distillation from a previous model to the trained model on the new data is the primary building block in data-focused methods [36], [37], [38], [39]. Prior-focused methods mitigate forgetting by estimating a distribution over the model parameters used prior when learning from new data [40], [41], [42], [43], [44], [45]. Parameter isolation methods isolate parameters for specific tasks and can guarantee maximal stability by fixing the parameter subsets of previous tasks. For example, Mallya and Lazebnik [46] have proposed an approach that uses weight-based pruning techniques to free up redundant parameters across all layers of a deep network after it has been trained for a task. Another approach built upon ideas from fixed network quantization and pruning is introduced in [47]. A different approach for continual learning is proposed in [48], namely, it searches for the best neural architecture for each coming task via sophisticated reinforcement learning strategies. The studies in [49], [50] also fall into the category of dynamic architecture-based continual learning solutions.

Most existing continual learning approaches are designed in a supervised fashion assuming all data from new tasks have been manually annotated. However, in many real-world applications of continual learning, e.g., learning from sensor data streams to make real-time classification, the availability of relevant labeled data is often low or even non-existing [51], [52], [53]. Most real-world data is usually not consistently labeled, i.e., there is no explicit indication of the exact periods of relevant events and occurrences of interesting trends, which breaks down the traditional supervised learning paradigm. Data labeling is mostly done manually by human experts. This process is, however, labor-intensive, time-consuming, and very expensive. Unsupervised continual learning, which is expected to tackle the aforementioned issues, has not been well studied [54]. Caron et al. [55] have proposed to iteratively cluster features and update the model with subsequently assigned pseudo labels obtained by applying a standard clustering algorithm. Another recent work proposes to perform clustering and model updates simultaneously to address the model's instability during the training phase [56]. However, these methods only work on static datasets and cannot learn new knowledge incrementally. In [57], the authors introduced a simple and effective method that, in an unsupervised setting, can be adapted to existing supervised continual learning approaches. The authors propose to use a pseudo label instead of the ground truth to make continual learning feasible in unsupervised mode. The pseudo labels of new data are obtained by applying a global clustering algorithm.

Evolving clustering models are good candidates to tackle concept drift scenarios. They have been designed to mine massive datasets or online continuous data streams in an unsupervised learning context by grouping and by summarizing data in a fast-incremental manner. Evolving clustering methods can process data stepwise and update and evolve cluster partitions in incremental learning steps [58], [59]. According to [58], different phases of an evolutionary clustering algorithm can be categorized into matching, accommodating new data, and model refinement. Dynamic clustering is also a form of online/incremental unsupervised learning [4], [6], [7], [9], [60]. However, it considers the incremental fashion of building the clustering model and self-adaptation of the built model. Dynamic clustering algorithms can split or merge the clusters based on the need.

Challenges: Data is often collected from unreliable sources, possibly having missing values and inaccurate labels. Hence, there is a need for a conceptually new learning framework to support continual and evolving learning under uncertainty and noise [51], [53]. In general, to take advantage of new developments in AI research, such as shared and continual learning [61], we need novel data mining and learning models. Those models should be capable of dealing with unlabeled data having large-scale missing and inaccurate labels, enabling learning from multiple data sources via distributed training and continual evolution of the model [62], [63] while efficiently

TABLE II
MAIN CATEGORIES OF SUPERVISED CONTINUAL LEARNING METHODS ACCORDING TO THE SURVEY PUBLISHED IN [20].

Method's category	**Sub-categories**	**Studies**	**Pros & Cons**
Replay	Rehearsal	[26], [27], [28], [29]	Limited scalability, Privacy issues, No clear policy for unbalanced tasks, Task-agnostic
	Pseudo-rehearsal	[30], [31], [32], [33]	
	Constrained	[34], [35]	
Regularization-based	Prior-focused	[40], [41], [42], [43], [44], [45]	Prioritizing privacy, Alleviated memory requirements, Task-agnostic
	Data-focused	[36], [37], [38], [39]	
Parameter isolation	Fixed Network	[47], [46]	Efficient memory, Prevents scalable class incremental setup
	Dynamic Architectures	[48], [49], [50]	

dealing with catastrophic forgetting and automatically adapting to real concept drift.

3) Federated learning: Federated learning (FL) has been introduced as promising collaborative learning, where edge devices such as smartphones, tablets, sensors, etc. keep their local data in their premises and exchange model parameters with a central server for global model aggregation [64], [65]. The global model is updated by averaging the local model parameters received by all the edge devices and is shared with them again. These operations are repeated at each iteration round. This setup has many advantages but also challenges such as expensive communication, systems heterogeneity due to the verity of devices in federated networks, and privacy concerns [66]. The iterative nature of FL requires massive communication between the central server and edge devices to train a global model [65]. The communication overhead at each iteration is not negligible, especially for complex models, large-scale applications, and high-frequency updates, and it becomes a challenge to be addressed [64], [65], [67]. Recently, many studies aiming to reduce communication costs have been proposed. For example, [68] use models of different sizes to address heterogeneous clients equipped with different computation and communication capabilities, while the work in [69] uses decentralized collaborative learning in combination with the master-slave model. The majority of solutions that address the problem of reducing network overhead in FL could be classified into two main categories. The first category incorporates works that reduce the total number of bits transferred for each local update through data compression. The second category includes studies that aim at reducing the number of local updates during the training process.

The authors of [70] propose an enhanced FL technique by introducing an asynchronous learning strategy on the clients and a temporally weighted aggregation of the local models on the server. The layers of the deep neural networks are categorized into shallow and deep layers. The parameters of the deep layers are updated less frequently than those of the shallow layers. In addition, a temporally weighted aggregation strategy is applied on the server to make use of the previously trained local models, thereby enhancing the accuracy and convergence of the central model. The paper [71] designs two novel strategies to reduce communication costs. The first relies on lossy compression on the global model sent server-to-client. The second strategy uses Federated Dropout (FD), which allows users to efficiently train locally on smaller subsets of the global model and reduces the client-to-server communication and the local computation. Deep Gradient Compression (DGC) is proposed to significantly reduce communication bandwidth [72]. The authors of [73] introduce a new compression framework, entitled Sparse Ternary Compression, that is specifically designed to meet the requirements of the FL environment. The authors of [74] implement a Federated Optimisation (FedOpt) approach by designing a novel compression algorithm for efficient communication. Then, they integrate additively homomorphic encryption with differential privacy to prevent data from being leaked. Malekijoo et al. [75] develop a novel framework that significantly decreases the size of updates while transferring weights from the deep learning model between clients and their servers. A novel algorithm, namely FetchSGD, that compresses model updates using a Count Sketch, and takes advantage of the mergeability of sketches to combine model updates from many workers, is proposed by [76]. Xu et al. [77] present a Federated Trained Ternary Quantization (FTTQ) algorithm, which optimizes the quantized networks on the clients through a self-learning quantization factor.

A novel FedMed method with adaptive aggregation is proposed using the topK strategy to select the top workers with the lowest losses to update the model parameters in each round in [74]. Asad et al. [78] have provided a novel filtering procedure on each local update that only transfers significant gradients to the server. The study proposed by [79] identifies the relevant updates of participants and uploads them only to the server. Specifically, at each round, the participants receive the global tendency and check the relevancy of their local updates with the global model. If they align, the updates are uploaded. An FL two-step client selection protocol based on resource constraints instead of the random client selection is proposed by [80]. FedPSO, a global model update algorithm, transmits the model weights only to the client that has provided the best score (such as accuracy or loss) to the cloud server [81].

Challenges: So far there is no evidence of how FL approaches are reducing the number of bits transferred compared to FL approaches that reduce the number of local updates. However, concerning the latter category of approaches, it is vital to find out more efficient FL schemes other than FedAvg, which converge with the same speed as FedAvg and apply to any FL applications [82]. For example, the studies in [83] and [84] have explored an approach that applies clustering optimization to bring efficiency and robustness in FL's communication: only the most representative updates are uploaded to the central server for reducing network communication costs.

IV. Preliminary Results

A. An inductive system monitoring approach for GNSS activation

In order to address the above challenges, we have designed a GNSS component activation model for mobile tracking devices which automatically detects indoor/outdoor environments using the radio signals received from LTE base stations [3]. We use an Inductive System Monitoring (ISM) technique [85] to model environmental scenarios captured by each tracker via extracting clusters of corresponding value ranges from base stations' signal strength. The ISM-based model is built by using the tracker's historical data labeled with GPS coordinates. The built model is further refined by applying it to the data without GPS location collected by the same device. This procedure allows us to identify the clusters that describe semi-outdoor scenarios. Thus, the model enables to discriminate between two outdoor environmental categories: open outdoor and semi-outdoor. Each cluster models an open outdoor or a semi-outdoor scenario by defining a range of allowable values for each base station in a given input vector. The vector of high values and the vector of low values in a cluster are considered as the cluster's representatives describing a specific environmental scenario. Evidently, the proposed model supplies the user with easily interpretable representations of the device's outdoor environmental scenarios. Note that the built ISM-based model does not contain the description of the indoor environmental scenarios, i.e., during the monitoring phase, data samples that do not fit any of the clusters are interpreted as belonging to the indoor environment. As a result, the built model is small and has modest requirements with respect to storage and computations.

B. Evaluation results

The proposed ISM-based GNSS activation approach is studied and evaluated on real-world data provided by Sony [3]. The used dataset contains radio signal measurements collected by five trackers and their geographical location in various environmental scenarios. We have explored the performance of the built ISM-based GNSS component activation model on this dataset in three different experiments. The obtained results have been analyzed and interesting patterns about the GNSS activation problem have been extracted. For example, we have conducted an experiment in which we use data with GPS coverage collected by each tracker to build a model representing the tracker's behavior. In addition, either the collected signals strengths by the trackers have been sorted in descending order, or they have been left as initially received. For testing, data without GPS coverage have been used. Table III lists the accuracy of the models for each tracker device with and without sorting the signal strengths. As one can observe, most models (M-di, $i = 1, \ldots, 5$) exhibit, except d5, higher accuracy in the case of sorted signal strengths. In addition, in another experiment, we have discovered that the models built on unshuffled data have shown higher performance. Furthermore, we have compared the performance of the model built on the data collected from all five devices with that of the individual trackers' models. The latter have demonstrated higher performance than the overall model. Evidently, the use of models with sorted and unshuffled signals is recommended. In addition, the customization of each tracker's model to the device specific environmental scenarios is preferred, since it ensures higher performance.

TABLE III
Model's accuracy (%) on data without GPS coverage for sorted and unsorted trackers' signal strengths

Model	Sorted signals	Unsorted signals
M-d1	99.89	64.44
M-d2	57.15	49.82
M-d3	61.49	60.40
M-d4	68.81	56.04
M-d5	71.94	72.83

The obtained evaluation results will be used for further improvement and optimization of the developed model (see our future plans in Section V). The company is currently evaluating and testing the model in the field.

V. Outlook

This paper describes the main objectives, identified challenges, and preliminary results of the DAISeN project. The main findings, valid for the reviewed research domains falling into the scope of DAISeN, reveal that in order to address the current challenges at the edge, we need novel resource and energy-efficient data mining algorithms and ML models robust to noisy, unlabeled, and missing data. Additionally, algorithms that enable learning from multiple data sources by distributed training and continual model adaptation are required.

In order to address the identified challenges, in the first half of the project, we have developed a novel GNSS component activation model for mobile tracking devices which is able to automatically detect indoor/outdoor environments based on the radio signals received from LTE base stations. The future research studies planned for the remaining period of the project involve the development of a domain integration GNSS activation technique that enables the integration of GNSS activation models built on different domains (devices/locations) into an overall model. In addition, we have the ambition to design a distributed GNSS activation framework that is enabled to create a shared model with the help of a large number of edge devices.

REFERENCES

[1] J. A. Manrique *et al.*, "Contrasting internet of things and wireless sensor network from a conceptual overview," in *2016 IEEE Int. Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data*, pp. 252–257. [Online]. Available: https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.66

[2] R. Arshad *et al.*, "Green iot: An investigation on energy saving practices for 2020 and beyond," *IEEE Access*, vol. 5, pp. 15 667–15 681, 2017. [Online]. Available: https://doi.org/10.1109/ACCESS.2017.2686092

[3] S. Abghari, V. Boeva, E. Casalicchio, and P. Exner, "An inductive system health monitoring approach for gnss activation," in *Artificial Intelligence Applications and Innovations*. Springer Nature Switzerland, 2022. [Online]. Available: https://doi.org/10.1007/978-3-031-08337-2_36

[4] V. M. Devagiri, V. Boeva, and E. Tsiporkova, "Split-merge evolutionary clustering for multi-view streaming data," *Procedia Computer Science*, vol. 176, pp. 460–469, 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.08.048

[5] C. Åleskog, V. M. Devagiri, and V. Boeva, *A Graph-Based Multi-view Clustering Approach for Continuous Pattern Mining*. Cham: Springer International Publishing, 2022, pp. 201–237. [Online]. Available: https://doi.org/10.1007/978-3-030-95239-6_8

[6] V. Boeva *et al.*, "Bipartite split-merge evolutionary clustering," in *Int. conference on agents and AI*. Springer, 2019, pp. 204–223. [Online]. Available: https://doi.org/10.1007/978-3-030-37494-5_11

[7] E. Lughofer, "A dynamic split-and-merge approach for evolving cluster models," *Evolving systems*, vol. 3, no. 3, pp. 135–151, 2012. [Online]. Available: https://doi.org/10.1007/s12530-012-9046-5

[8] C. Nordahl, V. Boeva, G. Håkan, and M. P. Netz, "Evolvecluster: an evolutionary clustering algorithm for streaming data," *Evolving Systems*. [Online]. Available: https://doi.org/10.1007/s12530-021-09408-y

[9] V. M. Devagiri, V. Boeva, and S. Abghari, "A multi-view clustering approach for analysis of streaming data," in *AI Applications and Innovations*, I. Maglogiannis, J. Macintyre, and L. Iliadis, Eds. Springer International Publishing, 2021, pp. 169–183. [Online]. Available: https://doi.org/10.1007/978-3-030-79150-6_14

[10] T. E. Bogale *et al.*, "Machine intelligence techniques for next-generation context-aware wireless networks," *Int. Telecommunication Union Journal*, 2018.

[11] Y. Zhu *et al.*, "A fast indoor/outdoor transition detection algorithm based on machine learning," *Sensors*, vol. 19, no. 4, p. 786, 2019. [Online]. Available: https://doi.org/10.3390/s19040786

[12] P. Bhargava *et al.*, "Senseme: a system for continuous, on-device, and multi-dimensional context and activity recognition," in *Proceedings of the 11th Int. Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2014, pp. 40–49. [Online]. Available: http://dx.doi.org/10.4108/icst.mobiquitous.2014.257654

[13] R. Sung *et al.*, "Sound based indoor and outdoor environment detection for seamless positioning handover," *ICT Express*, vol. 1, no. 3, pp. 106–109, 2015. [Online]. Available: https://doi.org/10.1016/j.icte.2016.02.001

[14] O. Canovas *et al.*, "Wifiboost: A terminal-based method for detection of indoor/outdoor places," in *Proceedings of the 11th Int. Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2014, pp. 352–353. [Online]. Available: https://doi.org/10.4108/icst.mobiquitous.2014.258063

[15] I. Ashraf *et al.*, "Magio: Magnetic field strength based indoor- outdoor detection with a commercial smartphone," *Micromachines*, vol. 9, no. 10, 2018. [Online]. Available: https://doi.org/10.3390/mi9100534

[16] W. Wang *et al.*, "Indoor-outdoor detection using a smart phone sensor," *Sensors*, vol. 16, no. 10, p. 1563, 2016. [Online]. Available: https://doi.org/10.3390/s16101563

[17] V. Radu *et al.*, "A semi-supervised learning approach for robust indoor-outdoor detection with smartphones," in *Proceedings of the 12th ACM Conf. on Embedded Network Sensor Systems*, 2014, p. 280–294. [Online]. Available: https://doi.org/10.1145/2668332.2668347

[18] T. Anagnostopoulos *et al.*, "Environmental exposure assessment using indoor/outdoor detection on smartphones," *Personal and Ubiquitous Computing*, vol. 21, no. 4, pp. 761–773, 2017. [Online]. Available: https://doi.org/10.1007/s00779-017-1028-y

[19] R. P. Souza *et al.*, "A big data-driven hybrid solution to the indoor-outdoor detection problem," *Big Data Research*, vol. 24, p. 100194, 2021. [Online]. Available: https://doi.org/10.1016/j.bdr.2021.100194

[20] M. D. Lange *et al.*, "A continual learning survey: Defying forgetting in classification tasks," *IEEE transactions on pattern analysis and machine intelligence*, vol. PP, 2021. [Online]. Available: https://doi.org/10.1109/TPAMI.2021.3057446

[21] R. M. French, "Catastrophic forgetting in connectionist networks," *Trends in cognitive sciences*, vol. 3, no. 4, pp. 128–135, 1999. [Online]. Available: https://doi.org/10.1016/S1364-6613(99)01294-2

[22] A. Gepperth and B. Hammer, "Incremental learning algorithms and applications," in *European symposium on artificial neural networks (ESANN)*, 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01418129

[23] P. Sprechmann *et al.*, "Memory-based parameter adaptation," in *Int. Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=rkfOvGbCW

[24] V. Moens and A. Zénon, "Learning and forgetting using reinforced bayesian change detection," *PLoS computational biology*, vol. 15, no. 4, p. e1006713, 2019. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1006713

[25] Y. Sun *et al.*, "Planning to be surprised: Optimal bayesian exploration in dynamic environments," in *Int. conf. on AGI*. Springer, 2011, pp. 41–51. [Online]. Available: https://doi.org/10.1007/978-3-642-22887-2_5

[26] M. De Lange and T. Tuytelaars, "Continual prototype evolution: Learning online from non-stationary data streams," in *Proc. of the IEEE/CVF Int. Conf. on Comp. Vision*, 2021, pp. 8250–8259. [Online]. Available: https://doi.org/10.1109/iccv48922.2021.00814

[27] S.-A. Rebuffi *et al.*, "icarl: Incremental classifier and representation learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2001–2010. [Online]. Available: https://doi.org/10.1109/CVPR.2017.587

[28] D. Isele and A. Cosgun, "Selective experience replay for lifelong learning," in *Proc. of the AAAI Conference on AI*, vol. 32, no. 1, 2018.

[29] D. Rolnick *et al.*, "Experience replay for continual learning," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[30] H. Shin *et al.*, "Continual learning with deep generative replay," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3294996.3295059

[31] F. Lavda *et al.*, "Continual classification learning using generative models," *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS) 2018*, 2018.

[32] J. Ramapuram *et al.*, "Lifelong generative modeling," *Neurocomputing*, vol. 404, pp. 381–400, 2020. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.02.115

[33] C. Atkinson *et al.*, "Pseudo-rehearsal: Achieving deep reinforcement learning without catastrophic forgetting," *Neurocomputing*, vol. 428, pp. 291–307, 2021. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.11.050

[34] D. Lopez-Paz and M. Ranzato, "Gradient episodic memory for continual learning," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3295222.3295393

[35] A. Chaudhry *et al.*, "Riemannian walk for incremental learning: Understanding forgetting and intransigence," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 532–547. [Online]. Available: https://doi.org/10.1007/978-3-030-01252-6_33

[36] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 12, pp. 2935–2947, 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2017.2773081

[37] H. Jung *et al.*, "Less-forgetting learning in deep neural networks," *Proceedings of the AAAI Conference on Artificial Intelligence*, no. 1, 2018. [Online]. Available: https://doi.org/10.1609/aaai.v32i1.11769

[38] A. Rannen *et al.*, "Encoder based lifelong learning," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1320–1328. [Online]. Available: https://doi.org/10.1109/ICCV.2017.148

[39] J. Zhang *et al.*, "Class-incremental learning via deep model consolidation," in *Proc. of the IEEE/CVF WACV*, 2020, pp. 1131–1140. [Online]. Available: https://doi.org/10.1109/WACV45572.2020.9093365

[40] J. Kirkpatrick *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the national academy of sciences*, vol. 114, no. 13, pp. 3521–3526, 2017. [Online]. Available: https://doi.org/10.1073/pnas.1611835114

[41] S.-W. Lee *et al.*, "Overcoming catastrophic forgetting by incremental moment matching," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3294996.3295218
[42] F. Zenke, B. Poole, and S. Ganguli, "Continual learning through synaptic intelligence," in *International Conference on Machine Learning*. PMLR, 2017, pp. 3987–3995.
[43] X. Liu *et al.*, "Rotate your networks: Better weight consolidation and less catastrophic forgetting," in *2018 24th ICPR*. IEEE, 2018, pp. 2262–2268. [Online]. Available: https://doi.org/10.1109/ICPR.2018.8545895
[44] R. Aljundi *et al.*, "Memory aware synapses: Learning what (not) to forget," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 139–154.
[45] A. Chaudhry *et al.*, "Efficient lifelong learning with a-GEM," in *Int. Conf. on Learning Representations*, 2019.
[46] A. Mallya and S. Lazebnik, "Packnet: Adding multiple tasks to a single network by iterative pruning," in *Proc. of IEEE CVPR*, 2018, pp. 7765–7773. [Online]. Available: https://doi.org/10.1109/CVPR.2018.00810
[47] A. Mallya *et al.*, "Piggyback: Adapting a single network to multiple tasks by learning to mask weights," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 67–82.
[48] J. Xu and Z. Zhu, "Reinforced continual learning," *Advances in Neural Information Processing Systems*, vol. 31, 2018. [Online]. Available: https://dl.acm.org/doi/10.5555/3326943.3327027
[49] R. Aljundi *et al.*, "Expert gate: Lifelong learning with a network of experts," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3366–3375. [Online]. Available: https://doi.org/10.1109/CVPR.2017.753
[50] A. Rosenfeld and J. K. Tsotsos, "Incremental learning through deep adaptation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 3, pp. 651–663, 2018. [Online]. Available: https://doi.org/10.1109/TPAMI.2018.2884462
[51] Z. Chen and B. Liu, "Lifelong machine learning," *Synthesis Lectures on AI and ML*, vol. 12, no. 3, pp. 1–207, 2018.
[52] G. De Francisci Morales *et al.*, "Iot big data stream mining," in *Proceedings of the 22nd ACM SIGKDD int. conference on knowledge discovery and data mining*, 2016, pp. 2119–2120. [Online]. Available: https://doi.org/10.1145/2939672.2945385
[53] H. M. Gomes *et al.*, "Machine learning for streaming data: state of the art, challenges, and opportunities," *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 2, pp. 6–22, 2019. [Online]. Available: https://doi.org/10.1145/3373464.3373470
[54] M. Masana *et al.*, "Class-incremental learning: survey and performance evaluation on image classification," *arXiv preprint arXiv:2010.15277*, 2020.
[55] M. Caron, Bojanowski *et al.*, "Deep clustering for unsupervised learning of visual features," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 132–149. [Online]. Available: https://doi.org/10.1007/978-3-030-01264-9_9
[56] X. Zhan *et al.*, "Online deep clustering for unsupervised representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 6688–6697. [Online]. Available: https://doi.org/10.1109/cvpr42600.2020.00672
[57] J. He and F. Zhu, "Unsupervised continual learning via pseudo labels," *arXiv preprint arXiv:2104.07164*, 2021.
[58] A. Bouchachia, "Evolving clustering: An asset for evolving systems," *IEEE SMC Newsletter*, vol. 36, pp. 1–6, 2011.
[59] M. Zopf *et al.*, "Sequential clustering and contextual importance measures for incremental update summarization," in *Proceedings of COLING 2016, the 26th Int. Conference on Computational Linguistics: Technical Papers*, 2016, pp. 1071–1082.
[60] M. Wang *et al.*, "A novel split-merge-evolve k clustering algorithm," in *2018 IEEE 4th Int. Conference on Big Data Computing Service and Applications (BigDataService)*. IEEE, 2018, pp. 229–236. [Online]. Available: https://doi.org/10.1109/BigDataService.2018.00041
[61] T. Mitchell *et al.*, "Never-ending learning," *Communications of the ACM*, vol. 61, no. 5, pp. 103–115, 2018. [Online]. Available: https://doi.org/10.1145/3191513
[62] I. Stoica *et al.*, "A berkeley view of systems challenges for ai," *arXiv preprint arXiv:1712.05855*, 2017.
[63] A. Tegen *et al.*, "Towards a taxonomy of interactive continual and multimodal learning for the internet of things," in *Adjunct Proceedings of the 2019 ACM Int. Joint Conference on Pervasive and Ubiquitous Computing and Int. Symposium on Wearable Computers*, 2019, pp. 524–528. [Online]. Available: https://doi.org/10.1145/3341162.3345603
[64] J. Konečnỳ *et al.*, "Federated learning: Strategies for improving communication efficiency," 2018.
[65] B. McMahan *et al.*, "Communication-efficient learning of deep networks from decentralized data," in *AI and statistics*. PMLR, 2017, pp. 1273–1282.
[66] T. Li *et al.*, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, pp. 50–60, 2020. [Online]. Available: https://doi.org/10.1109/MSP.2020.2975749
[67] W.-T. Chang and R. Tandon, "Communication efficient federated learning over multiple access channels," *arXiv preprint arXiv:2001.08737*, 2020.
[68] E. Diao *et al.*, "Heterofl: Computation and communication efficient federated learning for heterogeneous clients," in *Int. Conference on Learning Representations*, 2021.
[69] A. Reisizadeh *et al.*, "Fedpaq: A communication-efficient federated learning method with periodic averaging and quantization," in *Int. Conference on AI and Statistics*. PMLR, 2020, pp. 2021–2031.
[70] Y. Chen *et al.*, "Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 10, pp. 4229–4238, 2019. [Online]. Available: https://doi.org/10.1109/TNNLS.2019.2953131
[71] S. Caldas *et al.*, "Expanding the reach of federated learning by reducing client resource requirements," 2019. [Online]. Available: https://openreview.net/forum?id=SJlpM3RqKQ
[72] Y. Lin *et al.*, "Deep gradient compression: Reducing the communication bandwidth for distributed training," *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.
[73] F. Sattler *et al.*, "Robust and communication-efficient federated learning from non-iid data," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 9, pp. 3400–3413, 2019. [Online]. Available: https://doi.org/10.1109/TNNLS.2019.2944481
[74] X. Wu *et al.*, "Fedmed: A federated learning framework for language modeling," *Sensors*, vol. 20, no. 14, p. 4048, 2020. [Online]. Available: https://doi.org/10.3390/s20144048
[75] A. Malekijoo *et al.*, "Fedzip: A compression framework for communication-efficient federated learning," *arXiv preprint arXiv:2102.01593*, 2021.
[76] D. Rothchild *et al.*, "Fetchsgd: Communication-efficient federated learning with sketching," in *Int. Conference on Machine Learning*. PMLR, 2020, pp. 8253–8265.
[77] J. Xu *et al.*, "Ternary compression for communication-efficient federated learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2020. [Online]. Available: https://doi.org/10.1109/TNNLS.2020.3041185
[78] M. Asad *et al.*, "Ceep-fl: A comprehensive approach for communication efficiency and enhanced privacy in federated learning," *Applied Soft Computing*, vol. 104, p. 107235, 2021. [Online]. Available: https://doi.org/10.1016/j.asoc.2021.107235
[79] W. Luping *et al.*, "Cmfl: Mitigating communication overhead for federated learning," in *IEEE 39th international conference on distributed computing systems (ICDCS)*, 2019, pp. 954–964. [Online]. Available: https://doi.org/10.1109/ICDCS.2019.00099
[80] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *IEEE ICC*. IEEE, 2019, pp. 1–7. [Online]. Available: https://doi.org/10.1109/ICC.2019.8761315
[81] S. Park *et al.*, "Fedpso: federated learning using particle swarm optimization to reduce communication costs," *Sensors*, vol. 21, no. 2, p. 600, 2021. [Online]. Available: https://doi.org/10.3390/s21020600
[82] Q. Xia *et al.*, "A survey of federated learning for edge computing: Research problems and solutions," *High-Confidence Computing*, vol. 1, no. 1, p. 100008, 2021. [Online]. Available: https://doi.org/10.1016/j.hcc.2021.100008
[83] A. A. Al-Saedi, V. Boeva, and E. Casalicchio, "Reducing communication overhead of federated learning through clustering analysis," in *2021 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2021, pp. 1–7. [Online]. Available: https://doi.org/10.1109/ISCC53001.2021.9631391
[84] A. A. Al-Saedi, E. Casalicchio, and V. Boeva, "An energy-aware multi-criteria federated learning model for edge computing," in *2021 8th Int. Conf. on Future IoT and Cloud (FiCloud)*. IEEE, 2021, pp. 134–143. [Online]. Available: https://doi.org/10.1109/FiCloud49777.2021.00027
[85] D. L. Iverson, "Inductive system health monitoring," in *IC-AI*, 2004, pp. 605–611.