Artificial Intelligence Applications in industry @ PICUS Lab

Lorenzo De Donato, Antonino Ferraro, Antonio Galli, Stefano Marrone, Vincenzo Moscato, Elena Napoletano, Carlo Sansone, Stefania Santini, Giancarlo Sperlí, Valeria Vittorini

Department of Electrical Engineering and Information Technology (DIETI)
University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy
first.last@unina.it

Abstract

Artificial intelligence (AI) is being involved in a wide range of application fields, including critical ones. As a consequence of this wide spreading, AI-based systems are becoming more and more target of attacks aimed in circumventing them. Indeed, if on one hand industry is pushing toward a massive use of AI enhanced solutions, on the other hand it is not adequately supporting researches in end-to-end understating of capabilities and vulnerabilities of such systems. Since AI is very likely to be an important part of our everyday life in the very next future, it is crucial to build trust in AI systems. Although the solution is not straightforward, a crucial step in that direction is to raise awareness about security threats of these systems, from a technical perspective as well as from the governance and the ethical point of view. At the same time, AI can be used to enforce security in those scenarios where a continuous monitoring is needed for an early detection of dangerous activities. In this work, we briefly report the research activities on these topics we are working on at the Pattern analysis and Intelligent Computation for mUltimedia Systems (PICUS) Lab.

1 Introduction

The term "Artificial Intelligence" (or AI) has become more and more an integral part of the daily life of all of us. We are increasingly dealing with *smart* mobile phones, *intelligent* voice assistants, *robotic* chats, etc. Our interaction with these "intelligent systems" has become so predominant and widespread that even the world of industry has begun to use such AI in factory life, logistics and transportation. As a consequence, today's developing trend is increasingly focusing on AI based data-driven approaches in a wide range of application fields, including critical ones. At the same time, AI-based systems are becoming more and more target of attacks aimed in circumventing them. Indeed, if on one hand the industry is pushing toward a massive use of AI enhanced solutions, on the other hand it is not adequately supporting researches in end-to-end understating of capabilities and vulnerabilities of such systems. The results may be very (negatively) mediatic, especially when regarding critical scenarios.
In the last years, the pervasive use of Information and Communication Technologies (ICTs) in any industrial processes has significantly increased the amount of available data, whose analysis and processing improve the knowledge about the production processes, systems and equipment. This is leading the industrial world towards the so called ***Industry 4.0***, which mainly relies on the strong correlation between the physical and digital world of a production system enabling an efficient exchange of information [Esposito *et al.*, 2017; Galli *et al.*, 2020]. In particular, the analysis and processing of these data enables different drivers of analysis: i) reduce production costs; ii) obtain advantages in choosing decisional/operational strategies; iii) minimise failures, costs, repair times; iv) improve the safety of the human operator at work; v) optimise profit.
Since AI is very likely to be an important part of our everyday life in the very next future, it is crucial to build trust in AI systems. Although the solution is not straightforward, a crucial step in that direction is to raise awareness about security threats of these systems, from a technical perspective as well as from the governance and from the ethical point of view. At the same time, AI can be used to enforce security in those scenarios where a continuous monitoring is needed for an early detection of dangerous activities.
In the following, we briefly report the research activities on these topics we are working on at the Pattern analysis and Intelligent Computation for mUltimedia Systems (PICUSLab) Lab. These activities are intended to support, often by shedding lights on problems and issues, the realisation of a more trustworthy AI world.

2 AI for Industry: applications

In this section, we discuss about artificial intelligence activities developed by PICUSlab in Industry domain.

2.1 AI for the Rail Sector

In the last few years, there has been a growing interest in AI applications to railway systems. Such interest has been a consequence of the potential and opportunities enabled by AI powered solutions in combination with the other prominent technologies based on cloud computing, big data analytics and the Internet of Things. The AI challenge has been tackled by the European Union's Shift2Rail (S2R) programme with several research and innovation projects addressing aspects of digitalization, automation and optimization. In particular, a number of activities are currently being carried out at PICUSlab are in the context of the S2R project RAILS (Roadmaps for AI integration in the raiL Sector) [RAI, a], [RAI, b]. The overall objective of the RAILS research project is to investigate the potential of AI in the rail sector and contribute to the definition of roadmaps or fast uptake of AI and future research in the next generation signalling systems, operational intelligence, and network management. In the following, we briefly present the recent and ongoing works conducted at the laboratory as well as some preliminary results.

AI research in Railways and Challenges

Among the other activities, a taxonomy of AI to be used in the railway domain has been developed as well a comprehensive state-of-the-art of the available scientific literature and regulations [Bešinović *et al.*, 2021]. A survey involving the railway stakeholders has been conducted to identify the main challenges, milestones and obstacles related to the upcoming integration of AI in railway transportation. Fig. 1 illustrates the railway areas that have been addressed by the scientific literature on AI applications.

Some preliminary results can be derived from the analysis conducted so far. In particular, the outcomes of the survey are described in the Deliverable D1.3 of the RAILS project ([RAI, b]). The three most blocking factors for the adoption of AI in the rail sector are:

- safety, dependability and trustworthiness concerns;
- the lack of proper datasets for training the AI models;
- the lack of specific standards and regulations.

These results, together with the indications derived from the current railway research, suggest the importance of: a) defining pilot case studies/demonstrators to investigate the effects of AI solutions on safety-related applications, with the aim of producing the knowledge needed to drive standardisation; b) developing specific actions for railway datasets generation and sharing, including the definition of alliances/federations among relevant stakeholders; c) supporting the development of data-driven approaches aiming at enhancing safety of passengers and workers, such as data-driven risk assessment, accident prediction, avoidance and analysis (e.g., collisions, fire, accidents at level crossings, derailments, etc.).

A first step in the direction of collecting railway datasets has been done by surveying the publicly available datasets [Pappaterra *et al.*, 2021].

AI for rail safety and automation

The goal is to investigate the adoption of learning techniques and other AI methods for enhanced safety and rail automation. Specific activities aim at:

- studying the transferability of classical and novel approaches basing on the available results from rail and other transportation sectors. Toward this direction, a specific review work has been developed, see for example [Rajabli *et al.*, 2021], [Dirnfeld *et al.*, 2020].
- developing pilot case studies to investigate AI approaches and solutions in specific operational scenarios.

As to the second point, two representative case studies have been identified to inspect the effectiveness of AI-based methods for the deployment of next future autonomous trains: obstacle detection for intelligent train control and operation, and cooperative driving for virtual coupling of autonomous trains. These two case studies are meant to be representative of one of the main challenges arising when dealing with AI, i.e., the ability to tackle safety critical real time decisions exploiting limited data.

The high impact of the above mentioned case studies can be further appreciated when considering a general framework of the autonomous driving system architecture at the GAo4 Level, that can be abstracted from the automotive field. Here, the ability in performing obstacle detection and avoidance is one of the fundamental pillars of the perception layer, while virtual coupling is a critical issue to be solved at both planning and control layers.

AI for predictive maintenance and defect detection

The activity addresses AI techniques to support smart maintenance in railways. In particular, novel approaches are investigated to enable preventive condition-based maintenance by data analytics and machine learning, leveraging emerging technologies, such as the Industrial Internet of Things (IIoT) for sensing and actuation. Furthermore, advanced techniques based on Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) are addressed to support real time identification of anomalies through artificial vision enabled by smartcameras and other smart sensors. The Digital Twin paradigm is also exploited as a run time predictive model, based on the multi formalism integration and cross checking of diverse AI models.

Two pilot case studies have been selected to lead to *proofs of concept of AI applications and benchmarks*: Level Crossings (LC) and Digital Twins for railway stations maintenance. As to LC, the activities currently carried out aim at leveraging non-intrusive sensors for system monitoring and Remaining Useful Life (RUL) estimation, as well as investigating the usage of Transfer Learning to deal with limited availability of data.

Fig. 2 graphically describes the approach we adopted for audio signal recognition in the context of the wider LC case study: we re-used the same feature extractor of the VGGish [Hershey *et al.*, 2017] network, originally designed to extract 128-dimensional features from audio data [1], and built a new classifier for the detection of the LC warning bell signal.

The second case study concerns the usage of machine learning algorithms in a digital-twin based simulation model. The main task is to optimise the scheduling and planning of maintenance operations which could potentially increase the efficiency and robustness of the station system when abnormal event/unexpected congestion occurs, compared to the conventional or manual approaches.

In this phase, the results of the S2R IP3 research project named IN2SMART (Intelligent Innovative Smart Maintenance of Assets by integRated Technologies) are leveraged.

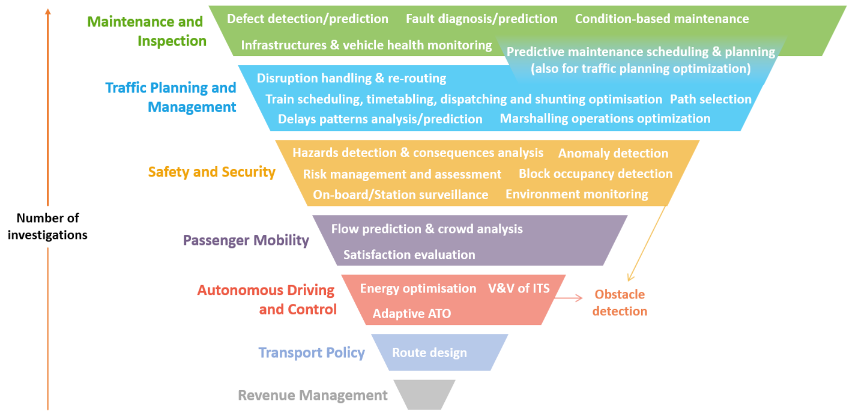

Figura 1: AI scientific publications in railways areas.

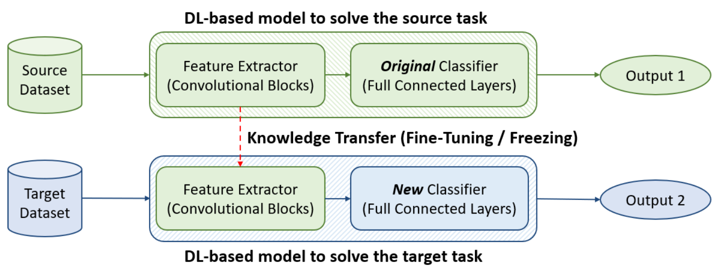

Figura 2: TL approach adopted for audio-signal recognition.

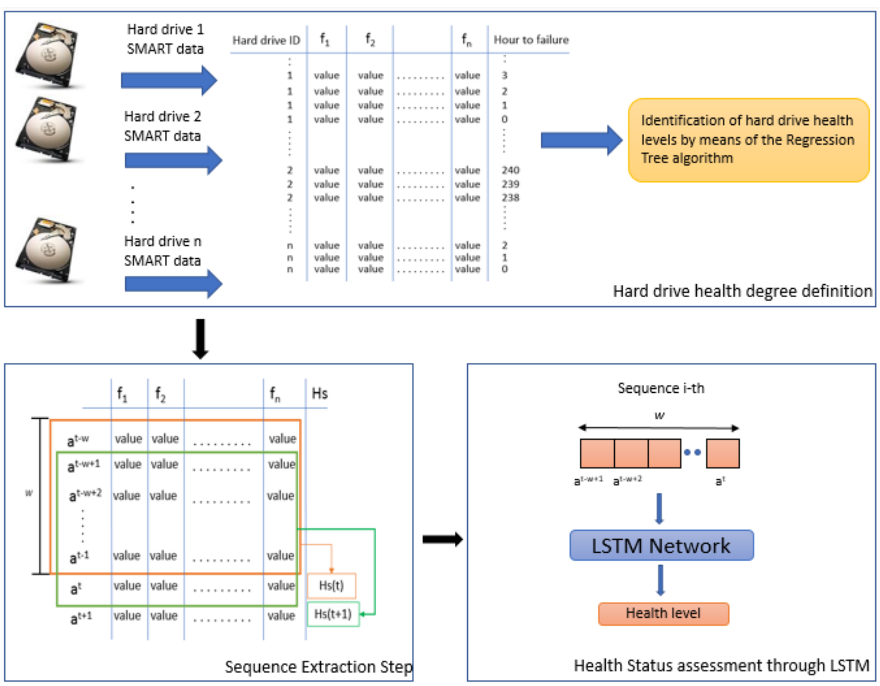

Figura 3: Workflow of the HDDs' level assessment.

2.2 Equipment Health assessment

The predictive maintenance is one of the main challenges in the industrial world aiming to jointly improve efficiency and operating times of the different production systems. In particular, predictive maintenance techniques based on machine and deep learning models have been designed to extend the Remaining Useful Life (RUL) of the equipment aiming to minimizing service shortage and data loss.

The increasing popularity of big data applications has made it so that storage systems are required to possess exabytes of capacity, usually resulting in millions of hard disk drives per data center. In particular, models based on *Self-Monitoring, Analysis and Reporting Technology* (SMART) have shown high accuracy levels by relying on internal attributes of HDDs as indicators of drive reliability. Importantly, most prediction systems analyze HDD failure as a binary classification task, simply distinguishing between good hard drives and those at high risk of failure. However, the complexity of the prediction task and the unbalanced nature of the data used in training have shown how these models' performance goes down significantly when they are tested on datasets representative of real-world environments [Aussel *et al.*, 2017].

For this reason, we propose a *Long Short Term Memory*-based framework [De Santo *et al.*, 2022], whose workflow has been shown in Figure 3, for assessing HDDs' health level, having the following characteristics:

- it automatically identifies the HDD health levels by considering the distribution of SMART attribute values over the time;
- it improves prediction accuracy by considering sequential dependencies in SMART attributes;
- it relies on an automated strategy for identifying the number and size of hard drive's health degree settings.

In [Galli *et al.*, 2020], we further show how a model-agnostic *explainable artificial intelligence* (XAI) tool (*SHapley Additive exPlanations* (SHAP) [Lundberg e Lee, 2017]) can be used to probe the results of our model, and support practitioners in their decision-making.

Riferimenti bibliografici

[Aussel *et al.*, 2017] Nicolas Aussel, Samuel Jaulin, Guillaume Gandon, Yohan Petetin, Eriza Fazli, e Sophie Chabridon. Predictive models of hard drive failures based on operational data. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 619–625. IEEE, 2017.

[Bešinović *et al.*, 2021] Nikola Bešinović, Lorenzo De Donato, Francesco Flammini, Rob M. P. Goverde, Zhiyuan Lin, Ronghui Liu, Stefano Marrone, Roberto Nardone, Tianli Tang, e Valeria Vittorini. Artificial intelligence in railway transport: Taxonomy, regulations and applications. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–14, 2021.

[De Santo *et al.*, 2022] Aniello De Santo, Antonio Galli, Michela Gravina, Vincenzo Moscato, e Giancarlo Sperlì. Deep learning for hdd health assessment: An application based on lstm. *IEEE Transactions on Computers*, 71(1):69–80, 2022.

[Dirnfeld *et al.*, 2020] Ruth Dirnfeld, Francesco Flammini, Stefano Marrone, Roberto Nardone, e Valeria Vittorini. Low-power wide-area networks in intelligent transportation: Review and opportunities for smart-railways. In *23rd IEEE International Conference on Intelligent Transportation Systems, ITSC 2020, Rhodes, Greece, September 20-23, 2020*, pages 1–7. IEEE, 2020.

[Esposito *et al.*, 2017] Christian Esposito, Massimo Ficco, Aniello Castiglione, Francesco Palmieri, e Huimin Lu. Loss-tolerant event communications within industrial internet of things by leveraging on game theoretic intelligence. *IEEE Internet of Things Journal*, 5(3):1679–1689, 2017.

[Galli *et al.*, 2020] Antonio Galli, Vincenzo Moscato, Giancarlo Sperlí, e Aniello De Santo. An explainable artificial intelligence methodology for hard disk fault prediction. In *International Conference on Database and Expert Systems Applications*, pages 403–413. Springer, 2020.

[Hershey *et al.*, 2017] Shawn Hershey, Sourish Chaudhuri, Daniel PW Ellis, Jort F Gemmeke, Aren Jansen, R Channing Moore, Manoj Plakal, Devin Platt, Rif A Saurous, Bryan Seybold, et al. CNN architectures for large-scale audio classification. In *Int. Conf. on Acoustics, Speech and Signal Processing*, pages 131–135. IEEE, 2017.

[Lundberg e Lee, 2017] Scott M Lundberg e Su-In Lee. A unified approach to interpreting model predictions. In *NIPS 30*, pages 4765–4774. 2017.

[Pappaterra *et al.*, 2021] Mauro José Pappaterra, Francesco Flammini, Valeria Vittorini, e Nikola Bešinović. A systematic review of artificial intelligence public datasets for railway applications. *Infrastructures*, 6(10), 2021.

[RAI, a] Roadmaps for AI integration in the raiL Sector (RAILS).

[RAI, b] Roadmaps for AI integration in the raiL Sector (RAILS).

[Rajabli *et al.*, 2021] Nijat Rajabli, Francesco Flammini, Roberto Nardone, e Valeria Vittorini. Software verification and validation of safe autonomous cars: A systematic literature review. *IEEE Access*, 9:4797–4819, 2021.

[1] https://research.google.com/audioset/download.html