

Advanced AI-based approaches in Industry 4.0 of the University of Naples Federico II node of the CINI-AIIS Lab

Lorenzo De Donato¹, Antonino Ferraro¹, Antonio Galli¹, Michela Gravina¹, Stefano Marrone¹, Vincenzo Moscato¹, Giancarlo Sperli¹, Valeria Vittorini¹ and Carlo Sansone^{1,*}

¹University of Naples, Federico II, Naples, Italy

Abstract

This paper provides an overview of recent advances in the application of artificial intelligence (AI) in industrial contexts such as data center and the railway industries carried out at the University of Naples Federico II node of the CINI-AIIS Lab. We discuss some challenges and opportunities associated with the adoption of AI in these industries. In data centers, AI is being used to optimize resource utilization, reduce energy consumption, and ensure high availability of services. Despite the potential benefits of AI, there are also challenges associated with its adoption, including the need for high-quality data, reliability and interpretability of AI-based systems, and ethical and legal concerns related to privacy, security, and bias. From the other hand, in the railway industry, AI is being used to optimize train schedules, reduce delays, support predictive maintenance operations and improve passenger safety by predicting and preventing accidents. In this paper, we focus on Hard Disk Drive health status assessment task and on an overview of AI techniques for railway domain in the context H2020 Shift2Rail project.

Keywords

Hard drive failure prediction, SMART, Long short-term memory, Industrial AI

1. Introduction

In recent years, there has been a growing interest in the application of Artificial Intelligence (AI) in industrial contexts such as manufacturing, transportation, and data centers [1, 2]. AI has the potential to transform these industries by enabling them to operate more efficiently, reduce costs, and improve safety and reliability. In particular, the railway and data center industries have witnessed significant growth in the application of AI.

Data centers, which are critical to the functioning of modern-day businesses, are facing challenges related to resource utilization, energy consumption, and service availability. The exponential growth in data traffic has led to an increase in energy consumption and carbon emissions, which has become a major concern for data center operators. AI-based techniques such as deep learning, reinforcement learning, and anomaly detection have

been used to optimize resource utilization, reduce energy consumption, and ensure high availability of services. These techniques can be used to predict and prevent hardware failures, optimize cooling systems, and allocate resources more efficiently.

The railway industry is facing numerous challenges in meeting the increasing demand for reliable and efficient transportation services. One of the main challenges is the need to optimize operational efficiency while ensuring passenger safety. AI-based techniques such as machine learning and predictive analytics have been used to analyze large volumes of data from various sources, including sensors, cameras, and social media, to gain insights into passenger behavior, track conditions, and train performance. This information can be used to optimize train schedules, reduce delays, and improve passenger safety by predicting and preventing accidents.

Despite the potential benefits of AI, there are also challenges associated with its adoption in industrial contexts. One of the main challenges is the need for high-quality data, which is essential for training AI models. In addition, there are concerns about the reliability and interpretability of AI-based systems, especially in safety-critical applications. Finally, there are also ethical and legal concerns related to the use of AI, including issues of privacy, security, and bias.

In this paper, we focus on Hard Disk Drive health status assessment task and on an overview of AI techniques for railway domain in the context H2020 Shift2Rail project.

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy

*Corresponding author.

✉ lorenzo.dedonato@unina.it (L. D. Donato);
antonino.ferraro@unina.it (A. Ferraro); antonio.galli@unina.it
(A. Galli); michela.gravina@unina.it (M. Gravina);
stefano.marrone@unina.it (S. Marrone);
vincenzo.moscato@unina.it (V. Moscato); giancarlo.sperli@unina.it
(G. Sperli); valeria.vittorini@unina.it (V. Vittorini);
carlo.sansone@unina.it (C. Sansone)

📞 0000-0001-9911-1517 (A. Galli); 0000-0001-5033-9617
(M. Gravina); 0000-0001-6852-0377 (S. Marrone);
0000-0002-0754-7696 (V. Moscato); 0000-0003-4033-3777 (G. Sperli);
0000-0002-8176-6950 (C. Sansone)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

2. HDD Health Status Assessment

Since hard drives often deteriorate gradually rather than abruptly, we argue that temporal analysis methods should be employed to model the sequential nature of the dependencies within SMART attributes over time. Thus, an approach to estimate the Remain Useful Life (RUL) of a HDD has been proposed, by automatically identifying specific health conditions on the basis of SMART attributes values. This methodology is grounded in three main steps:

- **Hard drive health degree definition:** in which status (or health level) is defined for each hard drive according to its time to failure;
- **Sequences extraction:** in which sequences in a specific time window are extracted for each hard drive;
- **Health Status assessment through LSTM:** in which a health level is associated to each temporal sequence.

In what follows, are described each component of the proposed framework in detail.

2.1. Health degree definition

The degradation of hard drives in real-world data centers is a gradual process that occurs over time. To account for this gradual decay, a system has been developed to measure the health status or level of a hard drive based on how much time is left before it fails. In contrast to the approach proposed in [3], this system includes an automated step for determining the health level of the hard drive.

More specifically, in this step has been considered only the hard drives that are going to fail, introducing for each of them an additional feature representing the *time before failure*. The data-set reports, for each hard disk, the temporally sorted sequence of SMART attributes with a specific sampling period. Denoting with m_j be the number of samples for the hard disk j , it is possible to associate each sample with an index i from 0 to $m_j - 1$, representing the number of samples that follow it in the sequence describing hard disk failure. As a consequence, the sample with index $i = 0$ is the last sample before failure. In Figure 1, *Time-to-failure* is the feature representing the time before failure for each hard drive whose meaning depends on the sampling period while f_1, f_2, \dots, f_n are the SMART attributes.

The idea is to build a Regression Tree (RT) for each SMART attribute f_i with $i = 1, 2, \dots, n$, having the feature representing the time before failure as a predictor and f_i as the numeric target value. Among all the resulting trees (one for each SMART attribute f_i), the one with the highest performance is selected, showing the attribute most

Hard drive ID	f_1	f_2	...	f_n	Time To Failure
...					...
1	value	value	value	3
1	value	value	value	2
1	value	value	value	1
1	value	value	value	0
...					...
2	value	value	value	240
2	value	value	value	239
2	value	value	value	238
...					...
n	value	value	value	2
n	value	value	value	1
n	value	value	value	0

Figure 1: *Time to failure* is a feature representing the time before failure for each hard drive sample, while f_1, f_2, \dots, f_n are the SMART attributes.

temporally dependent. Since the selected Regression Tree (RT) presents splits only on the feature *Time-to-failure*, the latter is used to distinguish hard drive health levels according to the time before failure.

A different level or status should be assigned to samples belonging to hard drives that will not fail since they have been excluded in this step. More specifically, the samples belonging to the hard drives that will not fail are labeled as *Good*.

2.2. Sequence extraction

To explore the temporal dependencies within the SMART features periodically collected for each hard drive, feature sequences in specific time windows (TW) have been extracted.

Let w and a^t be the time window size and the set of SMART features (f_1, f_2, \dots, f_n) at time t , respectively. Proposed model aims to predict hard drive health status at time $t + 1$ ($Hs(t + 1)$) considering the sequence $(a^{t-w+1}, \dots, a^{t-1}, a^t)$. For each a^t , the health status $Hs(t)$ is defined as the Regression Tree built, and the feature sequence for each hard drive at time t is extracted considering the $w - 1$ previous samples (cf. Figure 2). Each sequence results in a bidimensional array of size $w \times n$, where n is the number of SMART features considered. For each hard drive, sequences are extracted with a stride of one. It follows that $m_j - w + 1$ sequences are extracted for each hard drive, where m_j is the number of samples for the disk j .

For each sequence $(a^{t-w+1}, \dots, a^{t-1}, a^t)$, the hard drive's health level is defined by the health level of the set of features a^{t+1} . The result of this step is a sequence-based data-set — a set of bidimensional arrays, each associated to a health level representing the hard drive's health condition between two consecutive samples (i.e., a^t and a^{t+1}).

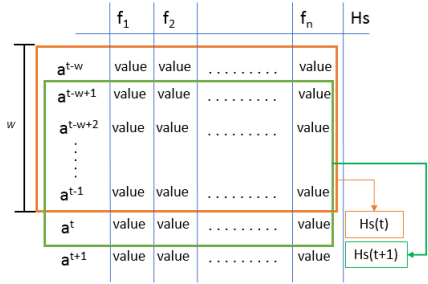


Figure 2: Sequence extraction step for a single hard drive.

2.3. Health Status assessment through LSTMs

Because of the sequential, gradually changing nature of the SMART features, it is important that designed model is able to capture dependencies across features over time. Long Short Term Memory networks (LSTMs) are extension to recurrent neural networks, explicitly designed with the purpose of learning long-term dependencies [4].

In the proposed framework, the input to each LSTM layer is a three-dimensional data structure of size $z \times w \times n$, where: z is the total number of sequences (or the batch size at each iteration); w is the size of each sequence — that size of a time window, in terms of time steps; n is the total number of features describing each time step.

Since the percentage of failed hard drives is often small compared to the percentage of good hard drives, the sequence extraction step may result in an unbalanced data-set with the majority of sequences belonging to the *Good* level. As a consequence, a data balancing step is introduced, so that the input to the network is a set of balanced data.

In particular, the sequence-based data-set is balanced by replicating the sequences belonging to the minority classes. Sequences replication is an efficient balancing strategy that avoids the polarization of the classification model on a single class without creating synthetic data or reducing the data-set size by sampling the instances belonging to the majority class. The implemented classification network has two stacked LSTM layers with 128 units, followed by a single dense layer.

2.4. Experimental Evaluation

The prediction performance of the model have been tested on the two different SMART data-set (Baidu and Backblaze)^{1,2}, and then compared with those of three other methods explored in the literature: a Classification

Tree model, a Random Forest model, and a model based on Multiclass Neural Networks.

2.4.1. Performance Evaluation

Since each sequence is associated with one of the levels presented in Section 2.1, the HDD's health level assessment as defined in this approach is a multiclass classification problem with multivariate input variables.

The performance of proposed approach is first evaluated in terms of accuracy, precision, and recall. Since the distinction between good and failed hard drives is preserved in the labelling of the data-set, we express the results in term of accuracy on good sequences (ACC_G) and accuracy on failed sequences (ACC_F) — respectively, the fraction of sequences correctly classified as *Good*, and the fraction of sequence classified as the health levels suggested by the regression trees. We also consider the evaluation criteria introduced in [3], and measure the accuracy of classifying good and failed sequences for a tolerance of misclassification up to one health level (ACC_G^{TOL} and ACC_F^{TOL}).

Finally, we evaluate performance in terms of failure prediction, by assessing *failure detection rate* (FDR) and *false alarm rate* (FAR) for each model. This is done by considering the levels *Good*, *Very Fair*, and *Fair* as *Hard drive good statuses*; and the levels *Soft Warning*, *Warning*, *Alert*, and *Red Alert* as *Hard drive failed statuses*.

2.5. Results and Discussions

We proposed a methodology to perform hard drive health status assessment exploiting the temporal dependencies of SMART attributes. In order to assess the effectiveness of the proposal, this section reports the performance of proposed methodology, and a comparison with several state-of-art approaches.

In particular, Table 1 and 2 show results of the LSTM based approach on the Baidu and Backblaze data-sets, respectively.

Performance is reported for different sizes of the time window (TW) used in the sequence extraction step. We explored time window sizes from 4 to 48 hours for the Baidu data-set, and from 5 to 15 days for the Backblaze data-set. For the latter, we considered a *prediction window* (q) varying from 15 to 45 days. As expected given the ability of LSTMs to learn long-distance dependencies, the best results are obtained with time windows of 48 hours and 15 days for the Baidu and Backblaze data-sets, respectively.

We compared our best results with respect to the sequence-independent models. The best results in terms of accuracy on failed sequences are obtained with RF for the Baidu data-set, and MNN for the Backblaze data-set

¹<http://pan.baidu.com/share/link?shareid=189977&uk=4278294944>

²<https://www.backblaze.com/b2/hard-drive-test-data.html>

TW SIZE [hour]	Accuracy	Precision	Recall	ACC_G	ACC_F	ACC_G^{TOL}	ACC_F^{TOL}	FDR	FAR
48	99.80%	99.1%	98.9%	99.83%	93.17%	99.89%	98.31%	98.2%	0.2%
36	98.78%	98.8%	98.7%	99.80%	91.89%	99.87%	97.45%	97.37%	0.2%
24	99.33%	98.9%	98.8%	99.66%	91.87%	99.74%	96.97%	97.64%	0.2%
12	98.71%	98.8%	98.6%	99.58%	78.06%	99.68%	90.54%	92.14%	0.4%
6	98.08%	98.3%	98.1%	99.43%	65.4%	99.59%	84.35%	86.8%	0.6%
4	97.74%	98.1%	97.8%	99.28%	60.29%	99.53%	82.47%	85.08%	0.6%

Table 1

Performance values for the LSTM models obtained by varying TW size on the Baidu data-set.

q [day]	TW SIZE [day]	Accuracy	Precision	Recall	ACC_G	ACC_F	ACC_G^{TOL}	ACC_F^{TOL}	FDR	FAR
15	5	95.88%	96.90%	95.10%	97.28%	66.56%	97.89%	98.08%	75.53%	2.82%
15	7	95.81%	97.10%	96.00%	97.02%	70.27%	97.93%	98.45%	79.34%	2.70%
30	5	94.54%	96.50%	94.60%	96.38%	56.07%	97.68%	88.30%	76.03%	2.73%
30	7	93.93%	96.80%	94.40%	95.59%	59.15%	97.07%	89.37%	80.70%	3.29%
30	10	95.25%	97.40%	96.10%	96.84%	61.84%	97.59%	91.35%	85.48%	2.73%
45	5	94.45%	96.70%	94.93%	95.95%	66.16%	97.80%	90.67%	78.30%	2.50%
45	7	95.82%	97.00%	95.85%	97.28%	68.34%	98.12%	89.37%	77.75%	2.17%
45	10	96.56%	97.72%	96.82%	97.71%	75.08%	98.36%	93.30%	84.18%	1.83%
45	14	98.45%	98.33%	98.34%	99.21%	84.49%	99.40%	96.65%	91.48%	0.72%

Table 2

Performance values for the LSTM models obtained by varying prediction window (q) and TW size on the Backblaze data-set.

Author	Methods	ACC_G	ACC_F	ACC_G^{TOL}	ACC_F^{TOL}
Xu et al. [3]	Multiclass NN	99.19%	16.01%	99.40%	43.34%
Xu et al. [3]	CRF	99.57%	28.51%	99.59%	61.30%
Xu et al. [3]	RNN	99.73%	41.05%	99.93%	64.86%
Our Approach	LSTM	99.83%	93.17%	99.89%	98.31%

Table 3

Comparison of our best model (LSTM - 48h) on the Baidu data-set with previously proposed models on the hard drive health status assessment task

Author	Methods	FDR	FAR
Xu et al.[3]	Multiclass NN	83.21%	0.60%
Xu et al.[3]	CRF	85.50%	0.22%
Xu et al.[3]	RNN	87.79%	0.004%
Li et al.[5]	CT	95.49%	0.09%
Zhu et al.[6]	BP NN	94.62%	0.48%
Shen et al.[7]	RF	97.67%	0.017%
Our Approach	LSTM	98.20%	0.20%

Table 4

Comparison of our best model (LSTM - 48h) on the Baidu data-set with previously proposed models on the hard drive failure prediction task.

Author	Methods	Accuracy	Precision	Recall
Zhang et al.[8]	LPAT+All	92.6%	89.30%	88.70%
Sun et al.[9]	TCNN	—	75.00%	67.00%
Basak et al.[10]	LSTM	—	84.35%	72.00%
Our Approach	LSTM	98.45%	98.33%	98.34%

Table 5

Comparison of our best model (LSTM - TW = 14 days and q = 45 days) on the Backblaze data-set with previously proposed models on the hard drive health status assessment task.

(98.13% and 96.17%). Results show that a sequence dependent approach provides higher performance than a sequence independent methodology, since the former is able to capture the SMART attribute temporal dependen-

Author	Methods	FDR	FAR
Shen et al.[7]	RF	94.89%	0.44%
Xiao et al.[11]	ORF	98.08%	0.66%
Our Approach	LSTM	98.20%	0.20%

Table 6

Comparison of our best model (LSTM - TW = 14 days and q = 45 days) on the Backblaze data-set with previously proposed models on the hard drive failure prediction task.

cies.

Finally, a comparison between the proposed methodology and some other proposals in the literature has been performed, which had also been tested on the SMART data-set. Tables 3, 4, 5 and 6 compare obtained best results on the Baidu and Backblaze data-sets with different approaches for hard drive health status assessment and hard drive failure prediction tasks.

To summarize, proposed approach outperforms all these models in terms of accuracy on failed sequences, FDR, and FAR both for hard drive health status assessment and hard drive failure prediction tasks. Importantly, experimental results demonstrate that the proposed approach is feasible for HDD health status assessment task due to the pre-processing phase and the definition of a specific model (LSTM) relying on the temporal sequence. Furthermore, the proposed approach has also been applied to other contexts such as railway rolling stock equipment [12], and IoT scenarios [13] resulting in efficient and effective also in these fields.

3. RAILS: Roadmaps for AI Integration in the rail Sector

The main objective of the H2020 Shift2Rail (S2R) project RAILS (Roadmaps for AI Integration in the rail Sector) is to investigate the potential of Artificial Intelligence (AI) techniques in the rail sector to contribute to the definition of recommendations and roadmaps for their implementation in various railway applications with a specific emphasis on three main pillars: Safety and Automation, Predictive Maintenance and Defect Detection, and Traffic Planning and Management. To be specific, the RAILS project has been focusing on six Proofs-of-Concept (PoCs), two for each pillar, intending to explore AI techniques to eventually identify their limitations, possible opportunities they can introduce, and directions that should be taken for the fast take-up of AI in railways.

In the last years, there has been an increasing interest in automatic train driving which led to the delineation of some promising approaches that, whether based on AI or not, could help to migrate towards a higher Grade of Automation, up to GoA 3/4, of railway lines. One of the main issues that are challenging this migration is the fact that intrusions or obstacles could affect the safety of such autonomous systems. To mitigate this factor, also under the S2R programme, investigations have already been performed towards the realisation of multi-source, multi-technology, and multi-level systems (exploiting, for example, multiple sensors, drones, and components installed on-board the train and on the trackside) which could help to improve the environmental perception of running trains. However, these systems are as effective as they are complex and potentially expensive. Therefore, within the RAILS project, efforts have been oriented towards the investigation of a “light” and cost-effective solution that could potentially help to deal with the detection of any kind of obstacles (both known and unknown a-priori) laying on rail tracks by exploiting of data coming from a camera in combination with supervised and unsupervised AI-based Computer Vision models (specifically, Deep Autoencoders) to segment rails and, eventually, detect possible obstacles. In addition to that, another, quite innovative, PoC is being addressed which deals with the integration of AI techniques for the introduction of Virtual Coupling (VC), widely examined within the automotive sector, in the context of autonomous trains. The main idea is to investigate the possibility of allowing two or multiple trains to act as a single convoy by exploiting Train-to-Train communication network, to reduce the headway between them, and a Reinforcement Learning approach to implement the VC control strategy. Interestingly, since one of the main challenges in the rail sector is the lack of suitable datasets to train AI algorithms, for these PoCs two simulators were built in order to col-

lect data for algorithms’ training and/or validate their effectiveness. Data Augmentation and Transfer Learning techniques were then (especially for the first PoC) used to increase the dimension of the dataset and improve AI algorithms performances respectively.

Concerning Predictive Maintenance and Defect Detection, attention is being posed to two of the most safety-critical railway assets. Investigations are being performed towards non-intrusive, cost-effective, and continuous monitoring of Level Crossings (LCs) by exploiting cameras and microphones in combination with Deep Learning approaches such as Convolutional Neural Networks and Object Detectors. The scope of this PoC encompasses two main directions: the first is to define a multi-modular approach that could support the identification of the health status and the estimation of the Remaining Useful Life of LC components; the second, on the other hand, is to explore techniques that can help to cope with the issue of the lack of available data, indeed, also in this case, one of the main challenges was to build suitable datasets for the effective training of AI algorithms and online repositories, real-life simulators, and Data Augmentation and Transfer Learning techniques were leveraged for this purpose. In addition to that, analyses are being carried out to define the contribution AI can bring to continuous monitoring and predictive maintenance of rolling stock components. Lastly, in the context of Traffic Planning and Management, the project is focusing on modelling and predicting train delays, a topic that is of primary importance when it comes to understanding the possible congestion of rail networks and, potentially, taking actions to mitigate this phenomenon. In this context, efforts have been posed on primary delays (i.e., delays generated by accidents/intrusions on rail tracks) prediction and estimation by means of a Graph Embedding Approach, while Graph Neural Networks were investigated to understand how accidents could trigger minor disturbances that, in turn, cause delays which propagate throughout the railway network and will cause, in the end, reactionary delays (i.e., knock-on or secondary delays).

To conclude, it is important to underline that the primary purpose of the aforementioned PoCs was to gather knowledge and evidence about the possible integration of AI in railways by exploring different techniques that could help to overcome possible issues (e.g., the problem of the lack of data). Hence, the final goal of the project would be to contribute to the definition of valuable recommendations and roadmaps, to be disclosed in the near future, that could support future research on AI in railways.

References

- [1] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, J. Barata, Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook, *IEEE Access* 8 (2020) 220121–220139.
- [2] M. Javaid, A. Haleem, R. P. Singh, R. Suman, Artificial intelligence applications for industry 4.0: A literature-based study, *Journal of Industrial Integration and Management* 7 (2022) 83–111.
- [3] C. Xu, G. Wang, X. Liu, D. Guo, T.-Y. Liu, Health status assessment and failure prediction for hard drives with recurrent neural networks, *IEEE Transactions on Computers* 65 (2016) 3502–3508.
- [4] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural computation* 9 (1997) 1735–1780.
- [5] J. Li, X. Ji, Y. Jia, B. Zhu, G. Wang, Z. Li, X. Liu, Hard drive failure prediction using classification and regression trees, in: 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, IEEE, 2014, pp. 383–394.
- [6] B. Zhu, G. Wang, X. Liu, D. Hu, S. Lin, J. Ma, Proactive drive failure prediction for large scale storage systems, in: 2013 IEEE 29th Symposium on Mass Storage Systems and Technologies (MSST), IEEE, 2013, pp. 1–5.
- [7] J. Shen, J. Wan, S.-J. Lim, L. Yu, Random-forest-based failure prediction for hard disk drives., *International Journal of Distributed Sensor Networks* 14 (2018).
- [8] J. Zhang, J. Wang, L. He, Z. Li, S. Y. Philip, Layerwise perturbation-based adversarial training for hard drive health degree prediction, in: 2018 IEEE International Conference on Data Mining (ICDM), IEEE, 2018, pp. 1428–1433.
- [9] X. Sun, K. Chakrabarty, R. Huang, Y. Chen, B. Zhao, H. Cao, Y. Han, X. Liang, L. Jiang, System-level hardware failure prediction using deep learning, in: 2019 56th ACM/IEEE Design Automation Conference (DAC), IEEE, 2019, pp. 1–6.
- [10] S. Basak, S. Sengupta, A. Dubey, Mechanisms for integrated feature normalization and remaining useful life estimation using lstms applied to hard-disks, in: 2019 IEEE International Conference on Smart Computing (SMARTCOMP), 2019, pp. 208–216. doi:10.1109/SMARTCOMP.2019.00055.
- [11] J. Xiao, Z. Xiong, S. Wu, Y. Yi, H. Jin, K. Hu, Disk failure prediction in data centers via online learning, in: Proceedings of the 47th International Conference on Parallel Processing, ACM, 2018, p. 35.
- [12] L. De Simone, E. Caputo, M. Cinque, A. Galli, V. Moscato, S. Russo, G. Cesaro, V. Criscuolo, G. Giannini, Lstm-based failure prediction for railway rolling stock equipment, *Expert Systems with Applications* (2023) 119767.
- [13] R. De Luca, A. Ferraro, A. Galli, M. Gallo, V. Moscato, G. Sperli, A deep attention based approach for predictive maintenance applications in iot scenarios, *Journal of Manufacturing Technology Management* (2023).