1.Introduction

* In the Industry 4.0 era, numerous sensors are being deployed across industries, generating data that can be used for Predictive Maintenance (PdM).
* Current PdM solutions are typically ad-hoc, designed for specific scenarios and unable to adapt to changes in setup, sensors, etc.
* The authors propose MachNet, a general, flexible Deep Learning (DL) architecture for PdM that can handle diverse sensor inputs, integrate prior information, and adapt to different problems.
* MachNet uses a sliding window of manufacturing processes instead of a temporal window, providing robustness and the ability to handle variable-length processes.

1. Related Work 2.1 Machine Learning based works

* Reviews various Machine Learning approaches to PdM, including Bayesian filters, ANNs, SVMs, RVMs, RNNs, etc.
* Highlights the need for hand-crafted feature engineering, which requires expertise and limits generalizability.

2.2 Deep Learning based proposals

* Discusses recent DL-based PdM methods using CNNs, LSTMs, autoencoders, etc.
* Notes their superior performance but lack of flexibility and need for fine-tuning for specific problems.
* Positions MachNet as an evolution of these architectures, with greater modularity and applicability.

1. Problem Definition

* Defines the PdM problem as estimating asset condition from heterogeneous data: sensor time-series (Arrays of Temporal Measurements or ATMs) and scalar parameters.
* Considers variable-length ATMs, unlike typical approaches that assume fixed duration.
* Discusses condition indicators (e.g., Remaining Useful Life, Health State) and the challenge of uncertain ground truth in real-world scenarios.

4. Network Architecture MachNet is a modular, two-stage deep learning architecture designed for Predictive Maintenance tasks. The first stage, Knowledge Extraction, consists of Temporal Modules (TMs) and Scalar Modules (SMs) that process raw sensor time-series and scalar parameters, respectively. The second stage, State Prediction, uses a Long Short-Term Memory (LSTM) network to fuse information from the TMs and SMs and estimate the asset's health state.

4.1 Window of processes MachNet introduces the concept of a sliding window of τ manufacturing processes, generalizing the typical temporal window approach. Each process in the window is represented by variable-length Arrays of Temporal Measurements (ATMs) from S sensors and R scalar parameters. The window is denoted as M^s\_(n,τ) for sensor s and P^r\_(n,τ) for scalar r, where n is the current process index.

4.2 Temporal-series modules (TMs) Each TM is a Convolutional Neural Network (CNN) that processes ATMs from a single sensor. The CNN architecture consists of two 1D convolutional layers followed by a Global Max Pooling (GMP) layer. The first convolutional layer has k\_conv1 kernel size and n\_conv1 channels, while the second layer uses multiple kernel sizes (k\_conv2) and n\_conv2 channels. The GMP layer adaptively summarizes the temporal information, producing a fixed-size output of dimension k\_G. This allows ATMs of variable length to be processed and combined into a fixed-size matrix.

4.3 Scalar modules (SMs) SMs are simple fully-connected (FC) layers that process scalar inputs. Each SM expands the dimension of the scalar parameters in the window to a fixed-size array of length k\_FC. The outputs of all SMs are combined into a matrix of the same dimensions as the TM outputs.

4.4 State predictor (SP) The SP is an LSTM network that fuses the outputs of the TMs and SMs to estimate the asset's health state. First, the TM and SM outputs are concatenated into a feature matrix F\_n. The LSTM, with hidden size h\_LSTM, processes F\_n to capture temporal dependencies across the τ processes in the window. The LSTM output is then passed through two FC layers (sizes h\_FC1 and h\_FC2) to produce the final health state estimate y\*\_n for the current process n.

5.1 Degradation model

* Describes the linear degradation models assumed for HS (SiMoDiM) and piecewise linear for RuL (C-MAPSS).

5.2 Performance evaluation

* Defines the evaluation metrics: Root Mean Squared Error (RMSE) and a RuL scoring function.

5.3 SiMoDiM: Coiler degradation

* Details the SiMoDiM dataset of coiler drums in a steel factory, with noisy sensor data, weak ground truth, and variable-length processes.
* Compares MachNet's HS estimation performance (RMSE) against ML (SVM) and Bayesian (DBF) baselines, demonstrating superior accuracy.

5.4 C-MAPSS: Turbofan degradation

* Describes the C-MAPSS dataset of simulated turbofan engine runs with piecewise linear RuL.
* Benchmarks MachNet's RuL prediction against state-of-the-art DL approaches, showing comparable or better performance, especially on simpler scenarios.

1. Conclusions

* Summarizes MachNet's key advantages: modularity, flexibility, ability to handle variable-length processes and prior information.
* Highlights strong empirical performance on both real (SiMoDiM) and simulated (C-MAPSS) datasets for HS estimation and RuL prediction, respectively.
* Discusses future work on new prognostic problems, knowledge transfer, interpretability, and integration of advanced DL techniques.

Definitions:

1. Predictive Maintenance (PdM):
   1. A proactive maintenance strategy that leverages data from sensors and other sources to estimate the health state of industrial assets and schedule maintenance activities accordingly.
2. Array of Temporal Measurements (ATM):
   1. A time-series of sensor readings collected during a single manufacturing process. In MachNet, ATMs can have variable lengths.
3. Convolutional Neural Network (CNN):
   1. A type of deep learning architecture commonly used for processing grid-like data, such as images or time-series. CNNs use convolutional layers to learn local patterns and pooling layers to summarize information.
4. Long Short-Term Memory (LSTM):
   1. A type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data. LSTMs use gating mechanisms to control the flow of information over time.
5. Transfer Learning:
   1. The practice of leveraging knowledge gained from solving one problem to improve performance on a related problem. In MachNet, transfer learning could involve reusing trained TMs or SMs across different PdM tasks.
6. Cyber-Physical System (CPS):
   1. A system that integrates computational and physical components, enabling real-time monitoring, control, and optimization of physical processes.