Proc. of the 9th Intl. Conf. of Asian Society for Precision Engg. and Nanotechnology (ASPEN 2022)
15–18 November 2022, Singapore. Edited by Nai Mui Ling Sharon and A. Senthil Kumar
doi://10.3850/978-981-18-6021-8_OR-08-0094

# Few-Shot Learning for Prescriptive Maintenance

**Yang Zhao, Yajuan Sun, Linshan Jiang, Ryan Chan Wei Yang**
Singapore Institute of Manufacturing Technology (SIMTech) @ Fusionopolis 2
2 Fusionopolis Way #08-04, Innovis Singapore 138634
# Corresponding Author / Email: zhao_yang@simtech.a-star.edu.sg, TEL: +65-86917869

KEYWORDS: Few-Shot Learning, Meta-Learning, Prescriptive Maintenance

*Machine learning has been applied to various areas, but one limitation of the industry to use machine learning al gorithms is that they do not have enough data. That is, most machine learning algorithms are still data-driven. Fe w-shot learning is a type of machine learning algorithm that focuses on processing the dataset containing few sam ples. Many different areas, such as medical, are exploring few-shot learning due to the sensitivity of their data. T he emergence of few-shot learning allows us to use the few data for data analytics.*

*Collecting data for the industry takes a long time because error cases are rare in daily operations. So, the indust ry takes one or two years to collect data, and then they can start data analytics. With few-shot learning, the time cycle can be shortened to some extent. The model-agnostic meta-learning (MAML) is one of the most popular fe w-shot learning algorithms widely used nowadays. Thus, we customize the MAML algorithm and apply it to our p rescriptive maintenance tasks to help improve the accuracy in predicting the failure when there are only hundreds of data samples. The experimental results on the real-world prescriptive maintenance dataset show that our custo mized algorithm achieves higher accuracy than the traditional algorithm.*

## 1. Introduction

In the Industry 4.0, industries are seeking opportunities to automate their production lines. With the increasing of machines, the maintenance of machines becomes a headache for many industries. Because they need to hire many people to monitor machines 24 hours a day in case of the breakdown of the machines, which does not effectively reduce the manpower. In addition, as the pandemic disease Covid-19 spreads the whole world, the labor-intensive manufacturing suffers a lot. Many workers have to stay at home to avoid the contact with other colleagues. Thus, to ensure the production line running smoothly, they have to reduce the maintenance staffs onsite. However, the maintenance cycle will be extended, which may cause the serious delay for the production line if some error happens unexpectedly. The necessity of automation cannot be ignored. Therefore, the prescriptive maintenance attracts quite some attentions these days, and more and more industries are willing to apply the prescriptive maintenance algorithms to their systems to improve the efficiency.

The prescriptive maintenance is that the system will predict the error happening and recommend required actions to fix the error [9]. The other problem faced by the industry is that data analysis for the prescriptive maintenance requires a large dataset, while most industries are not so modern that they can provide very few data or data contain privacy and they are unwilling to leak their data. Due to above two reasons, it is necessary to explore the prescriptive maintenance algorithm that can utilize the small dataset for the maintenance. Hereby, we recommend leveraging the few-shot learning algorithm.

Few-Shot Learning (FSL) targets at tasks which require limited examples [10]. The goal of few-shot meta-learning is to train a model that can quickly adapt to a new task using only a few datapoints and training iterations. It can rapidly generalize to new tasks of limited supervised experience by turning to prior knowledge, which mimics human's ability to acquire knowledge from few examples through generalization and analogy. It has been seen as a testbed for real artificial intelligence, a way to reduce the cost of those computationally expensive applications such as one-shot architecture search. And when the models and algorithms succeed for FSL, they naturally can apply for datasets of many-shots which are easier to learn.

In this work, we propose a meta-learning algorithm that is general and model-agnostic, in the sense that it can be directly applied to any learning problem and model that is trained with a gradient descent procedure. The key idea underlying our model is to train the model's initial parameters such that the model has maximal performance on a new task after the parameters have been updated through one or more gradient steps computed with a small amount of data from that new task. For the first stage of prescriptive maintenance, it is the anomaly detection. In this paper, we show that how we use few-shot learning solution to improve the accuracy when predicting failures in the first stage of the prescriptive maintenance.

## 2. Experiment

### 2.1 Problem Formulation

Unplanned machine downtime is very costly for industries, because of the idle production labor, spares value, etc. [4]. Anomaly detection is the first step for prescriptive maintenance. From anomaly detection, industries can predict the imminent failure, which may help to prevent the coming threatens. For industries, the unplanned machine downtime may result in a huge waste, because the repair and restart the large machine takes a long time and human resources are wasted during the period. Normally, it takes time to collect and label the large dataset. The traditional time cycle for the data analytics is quite long. So, the main challenge is to optimize the accuracy of the anomaly detection with few data.

### 2.2 Proposed Solution

We resolve the above challenge using a few-shot learning algorithm the model-agnostic meta-learning (MAML) [5]. MAML is designed to train the model's initial parameters, so that the model can achieve the satisfying performance on a new task with a small amount of data. MAML is applicable to various learning tasks, including classification, regression, etc. MAML aims to optimize parameters of the model, so that the model can achieve the maximal performance on a new task trained with a small amount data from the new task with few gradient updates [5]. Therefore, the model is easy to fine-tune when the algorithm is applied to a new dataset.

If we have trained a set of initialization parameters, it can be used to multiple new tasks across different industries who only need to provide few data. Specifically, MAML algorithm contains two stages:

(1) *Outer Loop is used to train the meta framework*. It learns a set of initialization parameters that can be used across different tasks.

(2) *Inner Loop is used to train every task*. Each task contains support set, validation set, and query set, which are used for training, validating, and testing every task. The inner loop helps to customize the initialization parameters to a specific task.

Thus, in the end, MAML will learn parameters for any task that samples a gradient step to obtain a satisfying result. When the large dataset for industries is infeasible, MAML has the potential to help save both industries and engineers' time and cost and achieve a good performance with few data.

### 2.2 Dataset

In our experiment, we use a real-world maintenance dataset - Robot Execution Failures Data Set [6]. It contains force and torque measurements on a robot after failure detection. They characterized each failure using 15 force/torque sample, which were collected immediately after failure detection at regular time intervals. The dataset includes 5 sub-datasets and 90 features. In our experiment, we use the LP1 sub-dataset, and its class distribution is as follows:

| Dataset | Number of Data | Class distribution |
|---|---|---|
| LP1: failures in approach to grasp position | 88 | 24% normal<br>19% collision<br>18% front collision<br>39% obstruction |

### 2.4 Model

Since every input data sample is one dimension, we design a 3-layer MLP classifier as input data sample is one dimension. The numbers of ReLUs in the three layers are 90,15 and 4 res pectively. A softmax function is used lastly to make the final pr ediction decision. Dropout is used during training to suppress ov erfitting. The loss function is cross-entropy loss function. The le arning rate for the inner loop of MAML is 0.01, and the step s ize is 0.4. Before training, the classification tasks should be con structed. $N$ different classes will be selected, and each class we randomly select $K$ samples for training. Specifically, $K$-shot mea ns that the input will use $K$ training samples for every class, an d totally $NK$ samples will be used for the $N$-way classification t ask. In our experiment, it involves 3-ways with 1 to 5 shots. T he meta-gradient that is calculated using $NK$ samples, i.e., at mo st 15 samples. All models were trained for 1000 iterations on a NVIDIA 2080ti GPU.

### 2.5 Results

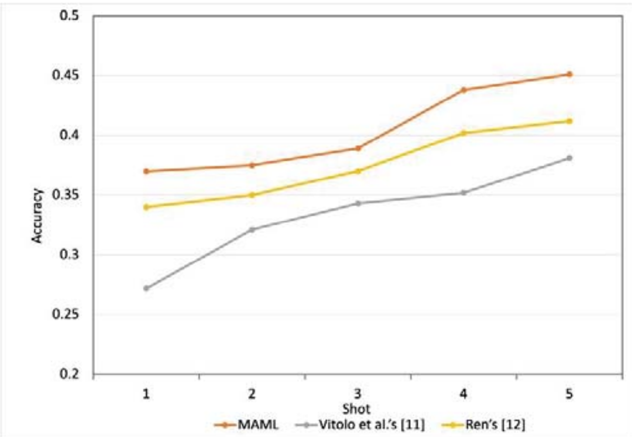
In our experiment, we randomly select three classes to construct the classification tasks. To evaluate the necessity of using few-shot learning, we evaluate the MAML algorithm together with Vitolo et al.'s solution [11] which is based on Convolutional Neural Networks (CNN) and Ren's solution [12] that is based on neighbor nearest algorithm. Following Vinyals et al.'s [7][8] design of the experiment, we use $N$-way classification with 1 or 5 shots. We present the experimental results in Table 1. With the existing amount of data, our proposed MLP based MAML algorithm outperforms other algorithms in accuracy. Compared with other algorithms which use the same number of data samples. It obviously that MAML is more efficient and effective.

*Table 1: Prediction accuracies when the number of shots is 1 or 5*

| | 3-ways 1-shot | 3-ways 5-shot |
|---|---|---|
| MAML | 0.37 +- 0.021 | 0.42 +- 0.031 |
| Vitolo et al.'s [11] | 0.35 +- 0.016 | 0.40 +- 0.012 |
| Ren's [12] | 0.27 +- 0.021 | 0.38 +- 0.022 |

Besides the above experiment, we also explore how the size of d ata affects the accuracy. The Figure 1 shows all the evaluation results for 1 to 5 shots under 3 ways. The results show that in each shot, the MAML can achieve better accuracy than other algorithms. Additional ly, from 1-shot to 5-shot, the accuracy increases slowly from 0.37 to 0.45 for MAML algorithm. The accuracies for classification usi ng other algorithms also increase with the number of samples in crease. Thus, the experimental results confirm that few-shot lear ning algorithm is more effective and efficient than other traditio nal data-driven algorithms when data are few. At the same time, the more samples the higher accuracy the learning algorithm ca n obtain, and this is applicable to all learning algorithms.

*Figure 1 : Prediction accuracies when the number of shots is 1 to 5*




## 3. Conclusions

We propose to leverage the few-shot learning （MLP based MAML framework）in the first stage of the prescriptive maintenance. The industry may leverage very little data to achieve satisfying accuracy. Thus, the few-shot learning-based prescriptive maintenance can help the industry save time and cost and leverage the few-shot learning in anomaly detection. The few-shot learning can also be used in the rest stages of prescriptive maintenance, for example, remaining useful life prediction and maintenance actions recommendation, and then it can be used to build a few-shot learning-based prescriptive maintenance framework in future work. The few-shot learning framework can be customized with various neural networks to help obtain good performance. Additionally, other commonly used few-shot learning algorithms are also worth exploring in the future.

## ACKNOWLEDGEMENT

This work is supported by A*STAR CAREER DEVELOPMENT FUND project - "Continuous Process Optimization and Prescriptive Quality Control for Electrochemical Process Using Machine Learning Approach".

## REFERENCES

[1] S. Jadon, "An overview of deep learning architectures in few-shot learning domain," arXiv preprint arXiv:2008.06365, 2020.L.

[2] M. Huisman, J. N. Van Rijn, and A. Plaat, "A survey of deep meta-

learning," Artificial Intelligence Review, vol. 54, no. 6, pp. 4483–4541,

2021.

[3] S. Ravi and H. Larochelle, "Optimization as a model for few-shot

learning," in ICLR, 2017.

[4] A. L. Patti, K. Watson, and J. H. Blackstone Jr, "The shape of protective capacity in unbalanced production systems with unplanned machine downtime," Production Planning and Control, vol. 19, no. 5, pp. 486–494, 2008.

[5] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in International Conference on Machine Learning. PMLR, 2017, pp. 1126–1135.

[6] P. M. Murphy and D. W. Aha, "UCI Repository of Machine Learning Databases [EB/OL]," http://archive.ics.uci.edu/ml/datasets.html.

[7] Liu, Xiaobo, Wei Teng, and Yibing Liu. "A Model-Agnostic Meta-Baseline Method for Few-Shot Fault Diagnosis of Wind Turbines." Sensors 22.9 (2022): 3288.

[8] Vinyals, Oriol, Blundell, Charles, Lillicrap, Tim, Wierstra, Daan, et al. Matching networks for one shot learning. In Neural Information Processing Systems (NIPS), 2016.

[9] Gordon, Christopher AK, and Efstratios N. Pistikopoulos. "Data-driven prescriptive maintenance toward fault- tolerant multiparametric control." AIChE Journal 68.6 (2022): e17489.

[10] Zhang, Jianwei, et al. "An Applicative Survey on Few-shot Learning." Recent Patents on Engineering 16.5 (2022): 104-124.

[11] Vitolo, Paola, et al. "Low-Power Detection and Classification for In-Sensor Predictive Maintenance Based on Vibration Monitoring." IEEE Sensors Journal 22.7 (2022): 6942-6951.

[12] Ren, Yali. "Optimizing predictive maintenance with machine learning for reliability improvement." ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg 7.3 (2021).