Automated Machine Learning architecture

For industrial IoT analytics

Introduction to SKF Enlight AI autonomous IIOT analytics

Machine Learning (ML) and Deep Learning (DL) are used in many fields, including computer vision, speech recognition and machine translation. Despite the increasing prevalence of these two methodologies, novice engineers still struggle to effectively apply these methodologies. At every step of Machine Learning, multiple decisions need to be made:

- Choosing between dozens of available ML and DL algorithms' preprocessing and cleaning methods.
- Tuning the hyperparameters of the selected approaches.
- Calibrating the number of trees or number of layers (and nodes at each layer) for the dataset

How effectively ML and DL are applied makes a significant difference and has a high impact on performance. The application of Machine Learning is tedious and time consuming, and requires the expertise of big data engineers and data scientists, rare commodities in today's labour market.

With increasing numbers of industrial plants looking to implement Machine Learning, there is a need for Machine Learning solutions that autonomously perform inference on a given dataset.

An Automated Machine and Deep Learning engine selects the optimal algorithm and hyperparameters in a data-driven way without any human intervention. In this article we describe how this is performed by the Enlight AI AutoML system.

Models selection, hyperparameters optimization and features extraction

A well-known method for optimizing Machine Learning hyperparameters is the Bayesian optimization, which iterates the following steps:

1. Build a probabilistic model to capture the relationship between hyperparameter settings and their performance.
2. Use the model to select optimal hyperparameter settings.
3. Run the selected Machine Learning model with those optimized hyperparameter settings.

This process can be generalized (to avoid overfitting) to jointly select algorithms, data preprocessing and data cleaning methods, and their hyperparameters as follows: the choices of classifier / predictor and preprocessing methods are top-level, categorical hyperparameters, and based on their settings the hyperparameters of the selected methods become active. The combined space can then be searched with Bayesian optimization methods that handle such high-dimensional, conditional spaces.

This Auto-MDL approach of using Bayesian optimization is used to automatically customize the optimal big data processing and unsupervised machine learning models to the appropriate industrial IOT analytics task.

The Enlight AI AutoML solution

AutoML is Enlight AI's implementation of the above idea. It contains a full Machine and Deep Learning pipeline which is responsible for missing values, categorical features, sparse and dense data, and rescaling/retreading /normalization of the data.

Next, the pipeline applies preprocessing and cleaning algorithms and an ML/DL algorithm.

AutoML includes dozens of ML algorithms, tens of preprocessing methods, and all their respective hyperparameters, yielding a total of hundreds of hyperparameters.

The optimizing performance in AutoML space of hundreds of hyperparameters can be slow. Enlight AI jumpstarts this process by using meta-learning to start the job from good hyperparameter settings originating from previous similar datasets (using a similarity function). When a new dataset is added, the algorithm looks for similar datasets as a starting point and applies the settings from the previous data set to the new one.

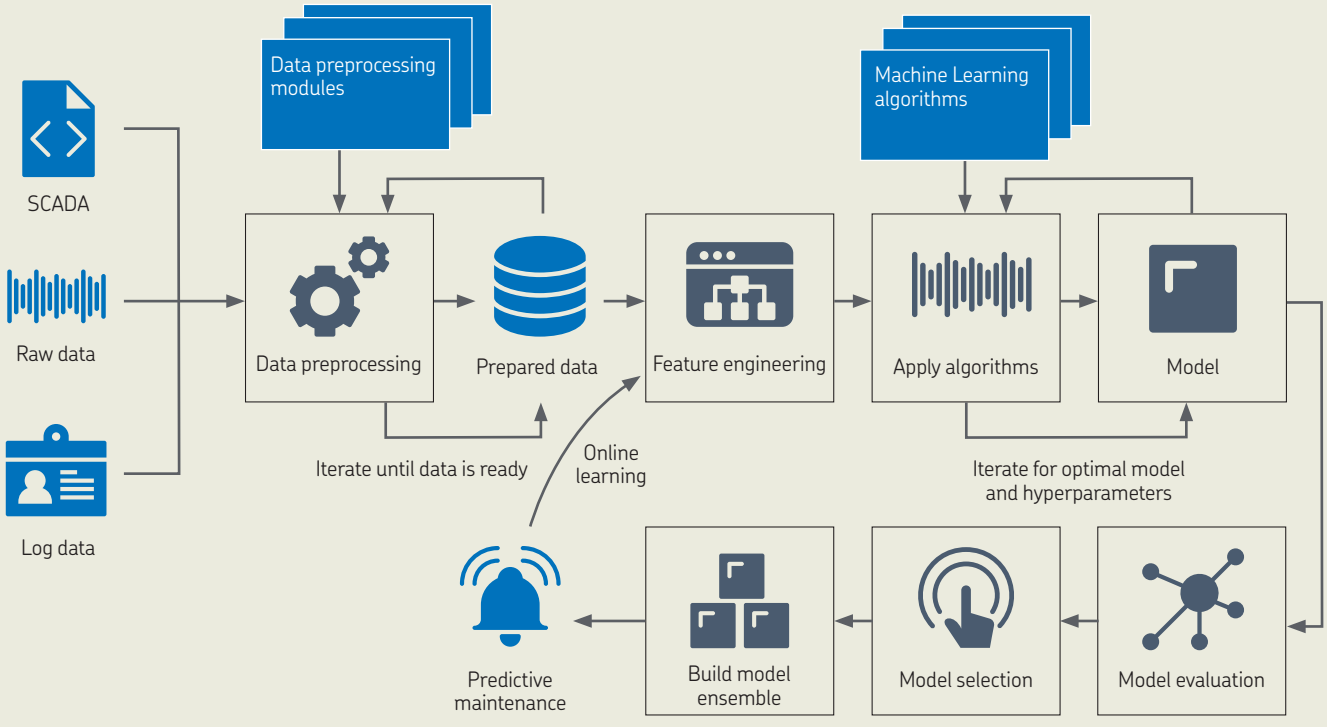
A second improvement was to automatically construct ensembles: instead of returning a single hyperparameter setting (as standard Bayesian optimization would), we automatically construct ensembles from the models trained during the Bayesian optimization. Specifically, we use model ensemble and stacking to create small, powerful ensembles with increased predictive power and robustness.

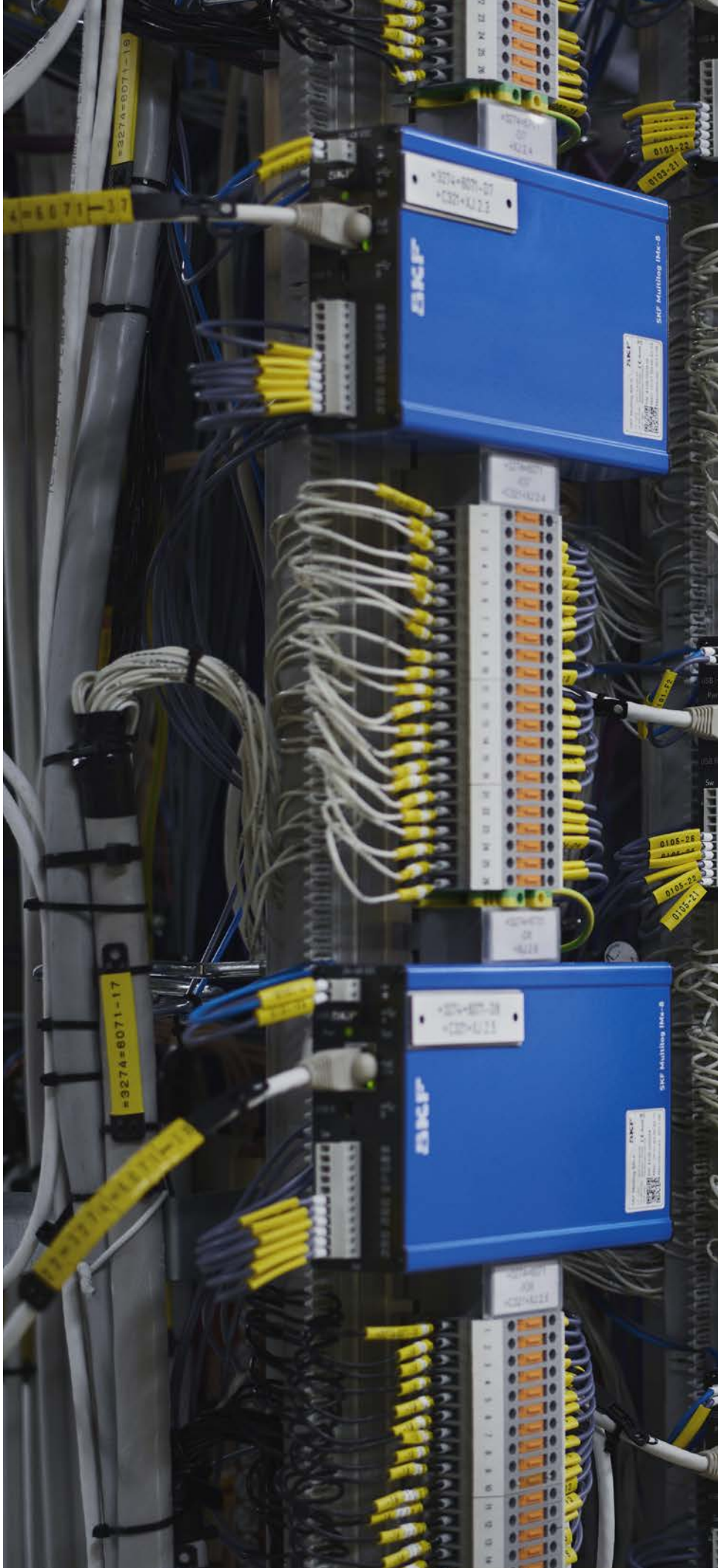
As more industrial plants migrate to the Smart Factory there is clearly a huge need for predictive asset maintenance.

The dearth of talented professionals with domain expertise in Deep Learning and Machine Learning will not be solved by increasing the supply of engineers and scientists. A technology solution is required so that facilities can adopt and scale predictive industrial IOT throughout a facility.

Enlight AI's AutoML is an off-the-shelf solution for industrial plants that are transitioning to Industry 4.0 and that want to invest a similar level of resources and capabilities as when they were operating in the older environment.

Hybrid analytics: Auto ML engine + SKF knowledge





skf.com/SKF-Enlight-AI

PUB SR/S9 18928 EN · July 2020