This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

An Automated Machine Learning Approach for Smart Waste Management Systems

David Rutqvist, Denis Kleyko, and Fredrik Blomstedt

***Abstract*—This article presents the use of automated machine learning for solving a practical problem of a real-life Smart Waste Management system. In particular, the focus of the article is on the problem of detection (i.e., binary classification) of an emptying of a recycling container using sensor measurements. Numerous data-driven methods for solving the problem were investigated in a realistic setting where most of the events were not actual emptyings. The investigated methods included the existing manually engineered model and its modification as well as conventional machines learning algorithms. The use of machine learning allowed improving the classification accuracy and recall of the existing manually engineered model from 86.8 % and 47.9 % to 99.1 % and 98.2 % respectively when using the best performing solution. This solution used a Random Forest classifier on a set of features based on the filling level at different given time spans. Finally, compared to the baseline existing manually engineered model, the best performing solution also improved the quality of forecasts for emptying time of recycling containers.**

Keywords*—smart waste management, emptying detection, classification algorithms, data mining, automated machine learning, grid search

I. Introduction

Machine learning is an area with a huge potential for the transformation of many areas of life and science including industrial informatics. In order to hasten the application of machine learning to real-world problems, the automated machine learning (AutoML) approach [1], [2] has been proposed. This article extends the AutoML approach with the data-driven methodology applied to industrial problems with existing (e.g., model-based) solutions. The methodology includes five steps:

- Collection of data, which can be used during the development and evaluation of solutions;
- The collected data are used to evaluate the existing solution to the problem;
- Parameters of the existing solution are optimised and evaluated based on the data;
- Conventional machine learning algorithms can be applied to the problem;
- The feature engineering methods are used to find if additional features could improve the results of the machine learning algorithms.

The methodology is applied to a problem within an area of waste management, which is one of the biggest challenges imposed by the rapid growth of the urban population. For example, in Europe each person is expected to yearly produce six tonnes of waste of materials used in the daily life [3]. An efficient strategy for facing the challenge of the waste management should address several directions including building a structured process for the waste disposal and maximising the recycling of the waste. When implementing these steps economical and environmental aspects should be taken into account. Waste transportation greatly affects both aspects and its optimisation can significantly increase the positive effects. At the same time, there is a clear requirement that in order to keep recycling stations clean they should be emptied at a right time. It is non-trivial to fulfil this requirement in a scenario with several hundreds of recycling stations (each with several containers) that are spread over a large geographical area.

A Smart Waste Management system implementing elements of Internet of Things is an enabling technology addressing the challenges of the waste transportation optimisation. It will allow each recycling container reporting its filling level. The advanced functionality of such a system will enable predicting the expected emptying time of a recycling container, i.e., the time when the container's filling level will achieve a certain critical value. Filling level predictions will allow avoiding redundant transportation without violating the overfilling requirement. However, the quality of filling level predictions will determine the efficiency of a Smart Waste Management system. There are several technical challenges for achieving a high quality predictions. Our analysis of an operating Smart Waste Management system revealed that one of these challenges is a problem of an accurate detection of a container being emptied using the measurements from a sensor mounted on top of a container. As it is demonstrated in Section II-B, the quality of filling level predictions depends on the correct detection of emptyings. Inaccurate detections devalue filling level predictions, therefore, detection of container emptyings is an integral step in obtaining qualitative predictions. Therefore, this article applies the proposed methodology to the problem of the emptying detection. Moreover, this study for the first time draws attention to the challenges and importance of the emptying detection for the functioning of Smart Waste Management systems.

A real-life Smart Waste Management system called the Smart Recycling ® is considered in this article. The article focuses exclusively on the emptying detection part of the considered Smart Waste Management system since the accurate emptying detection is a very important prerequisite for the accurate emptying time prediction. The article presents the

Manuscript received May 31, 2018; revised October 31, 2018; February 15, 2019; April 16, 2019; accepted May 3, 2019. Date of publication May 10, 2019; date of current version May 4, 2019.

During this DK has been supported by the GrowCPS project.

D. Rutqvist and F. Blomstedt are with the BnearIT AB, Luleå, Sweden.

*D. Kleyko is with the Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden, correspondence e-mail: denis.kleyko@ltu.se

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

existing solution to the emptying detection used in the system. It was manually engineered using the expert knowledge of the domain. It is also known that detection performance of the existing manually engineered model is a bottleneck for improving the emptying time predictions. Therefore, we propose using the data-driven approach for improving the quality of the emptying detection part of the system. It should be noted that the acquisition of a dataset used in this article was done as a part of the implementation work in this study.

The contribution of this article is twofold: conceptual and applied. The conceptual contribution is the methodology presented above. The applied contribution is the use of methodology for the development and investigation of several alternative solutions to the emptying detection problem. The investigated solutions included: 1) a modification of the existing manually engineered model as well as 2) the application of the conventional machine learning algorithms. The reported results indicate a significant improvement in the emptying detection performance. In particular, the accuracy increased from $86.8\ \%$ to $99.1\ \%$.

The article is structured as follows: Section II describes the outline of the considered Smart Waste Management system and formulates the problem of data-driven emptying detection. Section III briefly introduces the related work. Methods and data used in this article are described in a concise form in Section IV. The results of evaluation are presented and discussed in Section V. Section VI concludes the article.

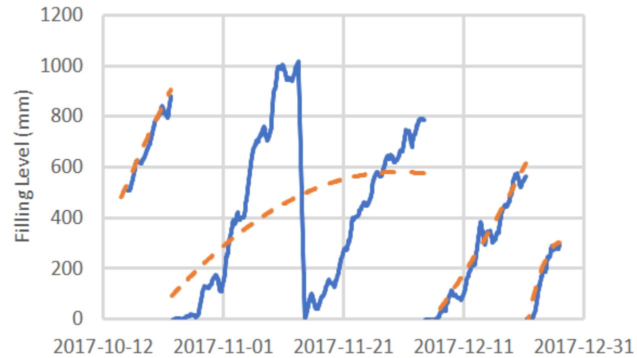


Fig. 1. Filling levels of a recycling container registered between October 13 and December 22. The solid line (blue) depict filling levels, which are a simple moving average of distance samples from the ultrasonic sensor. The dashed line (orange) is a polynomial function fitted between adjacent emptyings. A missed emptying and the effect it has on the function can be clearly seen around November 13. A detected emptying is depicted as an interruption in the solid line.

II. Problem formulation

A. Outline of the considered Smart Waste Management system

The considered Smart Waste Management system was created to address the challenge of efficient waste transportation. The system aims at optimising the waste management where a particular goal is to predict when a recycling container is going to be full. It is currently being deployed at every recycling station throughout the whole Sweden and the first $1{,}500$ recycling containers have been operating in the test mode for over a year in the southern part of Sweden. Each recycling container in the system is equipped with a sensor, which is mounted inside the container.

The sensors used in the system are a customised hardware solution. The hardware is equipped with an ultrasonic range sensor, an accelerometer, and a GSM module. The ultrasonic sensor measures a filling level of a recycling container regularly throughout the day at a configurable interval. Because the sensor is retrofitted to the ceiling of its container the measured range is a distance from the ceiling to the current waste level. In order to get an additional information for detecting potential emptyings, the hardware is also equipped with an accelerometer. The accelerometer continuously measures the acceleration of the corresponding container. If the acceleration exceeds a configurable threshold the sensor wakes the main processor with an interrupt. Then the processor measures how many interrupts it receives within a given time period. Thus, the accelerometer provides a *vibration strength* score. The obtained measurements are uploaded to a server using the GSM module. The processing of the sensory measurements for each recycling container is done on the server side. For example, the server transforms the distance to the waste level into the filling level in percentages. The server also collects the statistics and performs predictive analytics of the data received from the sensors. Finally, the server delivers the extracted information to a user interface.

B. Motivation for the accurate emptying detection in Smart Waste Management systems

Recall that the goal of a Smart Waste Management system is to predict emptying time, i.e., the time when a recycling container will be full enough to be emptied. In the considered system, it is defined that a recycling container should be emptied when its filling level reaches $90.0\ \%$. The statistics gathered from the live deployments have shown that the filling rate mostly follows either a line or a simple polynomial function. Therefore, the system can predict the filling level by fitting a regression model to the measurements reported by the ultrasonic sensor. The fitted regression model is used to extrapolate the filling level in the near future. Thus, given the regression model it is possible to estimate the emptying time.

A crucial prerequisite for this approach to function is that the regression model is built using only ultrasonic measurements obtained after the last emptying. Fig. 1 presents a counterexample of the fitted regression model (dashed line) in the case of a missed emptying. It is clear that the fitted regression model does not have any predictive power since it fitted the model to two filling cycles. Note also that the other regression models fitted to a single filling cycle closely followed the sensory measurements (solid line). Such situations as illustrated in Fig. 1 may lead to a container being overfilled, which, in turn, will result in decreased performance of the system. Therefore, this problem was identified as one of the key technical problems for the continued development of the considered Smart Waste Management system. The main measure, which should be taken to avoid such problematic situations, is the accurate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

detection of container emptyings. Moreover, it is desirable to achieve the accurate emptying detection without complicating the sensor, i.e., by using measurements only from the ultrasonic sensor and accelerometer.

C. Challenges in emptying detection

A simple solution to emptying detection would be a single threshold-based model where the values measured by either the ultrasonic sensor or accelerometer exceeded the threshold would lead to detection. There are, however, practical limitations for the use of such model. Due to the physical characteristics of the ultrasonic sensor, objects other than the actual waste level could be measured. For example, recycling containers usually have supporting structures or other parts related to the emptying mechanism, which interfere with ultrasonic pulses and, thus, create a false echo. Due to the false echo, the filling level will never reach zero even when the recycling container is actually empty. This fact invalidates the idea of measuring the absence of the waste for emptying detection. In the case of the accelerometer, since recycling containers are emptied by lifting them with a crane this event should trigger a single distinct vibration sample. However, in reality an emptying is not the only event, which generates vibration samples. Extra vibrations are often registered when waste is thrown in a recycling container. Thus, the use of accelerometer measurements could lead to many false detections.

The above arguments explain why it is not feasible to use a simple threshold model for the accurate detection of emptyings with either the ultrasonic range sensor or the accelerometer. However, it is possible to combine ultrasonic measurements and vibration strength scores in a more complex model with several thresholds. This idea was used to build the existing manually engineered model, which is presented in details in Section IV-B1.

III. Related work

AutoML [1], [2] is a recent research area and there were no extensive work applying it in industrial applications. Therefore, this section concentrates on the current state-of-the-art of the Smart Waste Management both in research and practical domains. The Smart Waste Management is very broad and includes many aspects. For the comprehensive survey of the area as an integral part of the Smart City concept readers are kindly referred to [4].

A. Commercial systems

Since the Smart Waste Management combines elements of Clean technology and Internet of Things as well as partially addresses the environmental issues, it is not surprising that there are several commercial systems competing in similar business niches.

The list of commercial systems can be divided into two types: retrofitted sensors and smart containers. Smart containers are specially made containers, e.g., urban dustbins or containers for cardboard measuring the filling level while mechanically compressing the cardboard. Such containers and dustbins are typically smaller in size and target restaurants, industries, and cities. Examples of commercial systems using smart containers are *Big Belly* [5] and *CleanCUBE* by Ecube [6].

Retrofitted sensors are of the same type as the considered Smart Recycling ® system and other examples of such systems are *Enevo* [7], *CleanFLEX* by Ecube [6], *Sensoneo* [8], *Onsense* [9], *Smart Waste* by Citibrain [10], and *Smart Bin* [11]. Amongst these systems, CleanFLEX is the most relevant to the Smart Recycling ® system since it also predicts filling levels of recycling containers.

B. Smart Waste Management research

The area of Smart Waste Management covers the whole lifecycle of a Smart Waste Management system. Since recycling containers are typically collected by trucks, areas such as fuel economy are also tightly related to the Smart Waste Management. Below the related work regarding the containers, their collection, and sensing techniques is presented. Other important aspects for design of such systems include: collecting data in an energy-efficient manner [12] where strategies can also be learned using data-driven techniques as well as questions related to a computing infrastructure [13] for supporting functionality of a system.

Currently, the research is often focused on the use of Internet of Things as an enabling technology for the Smart Waste Management. For example, several studies suggested using various technical implementations to help solving the problem of Smart Waste Management in large cities [14], [15], [16], [17], [18], [19]. In [14] a flexible and scalable platform was suggested for information sharing between heterogeneous devices. The goal of the platform was to monitor the level of a dustbin's fullness in order to avoid collecting semi-empty dustbins. The data could then be exported to other decision algorithms for determining the optimal number of trucks or dustbins in an area. In comparison to the Smart Recycling ® system, the platform targeted smaller residential dustbins while the considered system targets larger containers. In other words, there is a distinction between dustbins and containers where dustbins are most often used by residents and in public spaces in cities, and containers are larger (2 m^3 and more) waste receptacles commonly collected by large trucks equipped with a crane. Another work [15] goes through the need and use-cases for the Smart Waste Management and suggests a solution with ultrasonic sensors uploading their level to a cloud for further data processing, which strongly resembles the Smart Recycling ® system. However, the work only describes the system and its goals without considering emptying detection in details. A literature review of different analysis and optimisation techniques for urban solid waste management suggested that information technology will aid in designing a smart and green urban waste collection system [16]. The literature review identified both radio-frequency identification and sonar technologies as techniques of interest, both of which have been used or is currently used in the Smart Recycling ® system. Finally, two previous case-studies of the Smart Recycling ® system have been conducted in [20], [21] where [20] focused

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

on the impact the system made to its stakeholders while [21] described the development process behind the system. These case-studies, however, do not focus on the considered problem of the detection of emptying a recycling container since it is a rather specific topic arising in relation to the chosen sensing techniques. Nevertheless, it is important for the functioning of the whole system. As it was demonstrated in Fig. 1 for systems with retrofitted sensors, correct emptying detection is a key component for enabling accurate predictions of filling levels. Predictions of filling levels, in turn, affect the performance of the whole Smart Waste Management system since decisions are based on these predictions. Thus, solving the problem of emptying detection is a crucial task during the design phase of the system's life-cycle. To the best of our knowledge, this article is the first attempt to study solutions to this problem.

IV. Materials and Methods

A. Materials

1) Obtaining labelled data: In machine learning formulation, the considered problem of detecting a container emptying can be seen as a classification problem with two classes: emptying (positive) and non-emptying (negative) since the presence of vibration strength score identifies potential intervals with emptyings. Obtaining a dataset for this study was a non-trivial process since there is no obvious way to automatically label sensor measurements with correct class labels. Förpacknings- och Tidningsinsamlingen AB (FTI) is the packaging and newspaper collection service in Sweden. FTI has a system which allows hauliers registering when they have emptied a container, the approximate filling level at emptying, and the fraction emptied. It was possible to acquire this data from FTI for the same time period (approximately one year) as the actual available sensor measurements from the considered Smart Waste Management system.

The challenge, however, is that hauliers freely decide when to register emptyings and it is common that they register them several hours later. Thus, it was not feasible to simply match the time stamp provided by FTI with the corresponding sensor measurements (specifically with vibration strength scores) from the system. Therefore, a labelling system was developed. It holds a list of unlabelled time stamps that are of particular interest and a collection of time stamps labelled with FTI data. Out of the unlabelled time stamps a random point of interest can be picked and sensor measurements are loaded for a time span around that point. The user then manually selects a number of time stamps and labels them either as an emptying or non-emptying. This allows the user labelling several time stamps at once, e.g., if there is no actual emptying for the time span all samples for the time span can be labelled as non-emptyings. The FTI data were used to identify the actual emptyings as they are usually located within several hours from the time registered in FTI data. During the labelling process, particular attention was paid to the acquisition of the challenging time steps, i.e., emptyings, which were hard to detect, or non-emptyings, which could easily be misinterpreted as emptyings. The outcome from the labelling process was a list of time stamps with a container ID and a classification label, i.e., whether the time stamp was an emptying or not. This list was then used to extract the features from the corresponding sensor measurements.

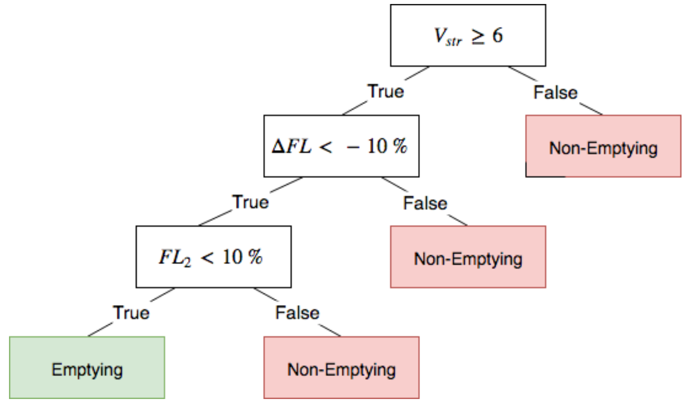


Fig. 2. Decision rules for the existing manually engineered model. The rules are traversed from the root until a class label (coloured) is assigned. Note the negative sign in the filling level change (ΔFL), hence, a decrease.

2) Balance and distribution of datasets: During the labelling process, approximately 27,000 time stamps were labelled. The labelled time stamps included different container types, i.e., types of recyclable materials such as glass, cardboard, and paper. The dataset contains a lot more non-emptyings than emptyings. The approximate ratios to the full size of the dataset are 77 % and 23 % respectively. This distribution resembles the real one since it is common to observe vibration strength scores without the actual emptying.

Finally, the whole dataset was split randomly into datasets for training and testing. The training dataset included 70 % of the whole dataset while the other 30 % composed the testing dataset. In particular, the testing dataset included 1901 emptyings and 6109 non-emptyings.

B. Methods

1) The existing manually engineered model: The existing manually engineered model to the emptying detection uses a set of static rules. These rules are applied every second hour to check for new vibration strength score (denoted as V_{str}). The existing manually engineered model uses only the vibration strength score and the ultrasonic measurements (filling level) directly preceding and succeeding the vibration strength score in time. This makes the rules susceptible to minor measurement errors in the near proximity of an emptying. The filling level before a vibration strength score is denoted as FL_1 and the filling level after is denoted as FL_2. The filling level change, ΔFL, is calculated as $\Delta FL = FL_2 - FL_1$. In order for a vibration strength score to be considered to be an emptying, V_{str} must exceed six, i.e., more than five interrupts must have been registered. Next, the filling level must drop more than 10 % (i.e., $\Delta FL < -10$ %) and the filling level after (i.e., FL_2) must be below 10 %. Only in this case, the emptying will be detected otherwise the event will be treated as a non-emptying. The rules can be specified in the form of a tree as illustrated in Fig. 2.

2) Conventional classification algorithms: The following six conventional classification algorithms were used to solve the problem[1]:

- Artificial Neural Network (ANN; [22], chapter 18.7);
- k-Nearest Neighbours (kNN; [22], chapter 18.8);
- Logistic Regression (LR; [22], chapter 18.6);
- Support Vector Machine (SVM; [22], chapter 18.9);
- Decision Tree (DT; [22], chapter 18.3);
- Random Forest (RF; [22], chapter 18.10).

3) Performance metrics: The main tool for assessing the performance of a model on the binary classification problem is a confusion matrix and it will be used below to present the results for some of the solutions. However, it is also convenient to use a single number (e.g., accuracy) for comparison reasons but since the dataset is skewed towards non-emptyings, the accuracy alone will not fully represent the classification performance. Therefore, several additional performance metrics [23] are considered:

- Recall;
- F1 score;
- Matthews Correlation Coefficient (MCC score).

V. Results

This section presents the results of nine solutions for the considered emptying detection problem. First, the performance of the existing and the optimised manually engineered models is presented. Second, the six conventional classification algorithms are evaluated using the same features as in the existing manually engineered model (i.e., V_{str}, FL_2, and ΔFL). Third, the extended set of features is considered. Finally, all obtained results are summarised and discussed.

A. The existing and the optimised manually engineered models

Recall that the existing manually engineered model consists of several combined rules and includes three thresholds. In the existing model these thresholds were set using the expert knowledge. An alternative way to configure the thresholds is via optimising the performance of the model on the training dataset. The optimisation was done via the grid search over the parameter space. The MCC score on the training dataset was evaluated for each combination. Since the MCC score varies between -1 and 1 it can be naturally mapped onto a heat map for the visualisation of the optimisation results. Fig. 3 presents such two-dimensional heat maps. One can draw intuitive conclusions from them, for example, that there is a distinct performance limit at $\Delta FL = 0$, which indicates that after the actual emptying, the filling level should not increase. Another observation is that FL_2 does not seem to be a very important parameter since the colours are consistent for a broad range of values.

The combination of values with the highest MCC determined the optimised threshold values. Table I contrasts the threshold values for the existing and optimised models. Unintuitively, the optimisation results suggest that the vibration strength threshold V_{str} should not be considered since it was assigned value zero, which is less than the minimum possible value of 1. Note however, that the vibration strength is still used for detecting time stamps to be classified because most of the time the vibration strength equals zero.

[1]It should be noted that we have not relied on any existing AutoML tool while performing the experiments.

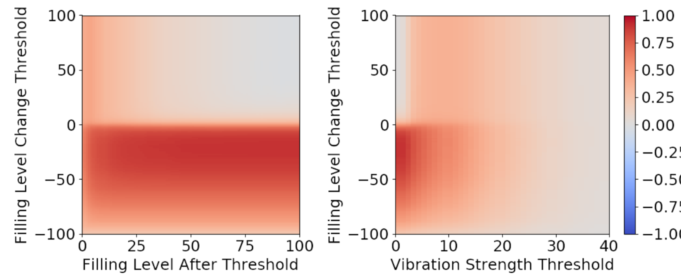


Fig. 3. Heat maps of the MCC score on the training dataset for different threshold values of the manually engineered model. Left panel: FL_2 against ΔFL; the vibration strength threshold V_{str} was kept fixed at zero. Right panel: V_{str} against ΔFL; the filling level after threshold FL_2 was kept fixed at 77 %. Both fixed values were the optimal ones according to the optimisation results.

TABLE I. Existing and optimised threshold values for the manually engineered model.

Parameter	Existing	Optimised
V_{str}	≥ 6	≥ 0
FL_2	< 10 %	< 77 %
ΔFL	< -10 %	< -20 %

Both the existing and the optimised manually engineered models were evaluated using the testing dataset. The corresponding confusion matrices are presented in Tables II and III.

TABLE II. Confusion matrix for the existing manually engineered model.

		Ground truth	
		Emptying	Non-emptying
Predicted	Emptying	911	68
	Non-emptying	990	6041

TABLE III. Confusion matrix for the optimised manually engineered model.

		Ground truth	
		Emptying	Non-emptying
Predicted	Emptying	1720	90
	Non-emptying	181	6091

The existing manually engineered model achieved rather high accuracy of 86.8 %. However, this should be attributed mostly to the accurate classification of non-emptyings since the obtained recall was only 47.9 %. The poor recall indicates the main performance drawback of the existing manually engineered model: it only correctly classified 911 out of 1901 actual emptyings. The optimised manually engineered model solved the poor recall problem to a large extend. It achieved 90.5 % recall and 96.6 % accuracy. With respect to the MCC score there was also a large improvement from 0.608 to 0.905. Thus, even a simple data-driven optimisation allowed achieving a huge performance improvement. Next, the performance of the conventional classification algorithms is evaluated.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

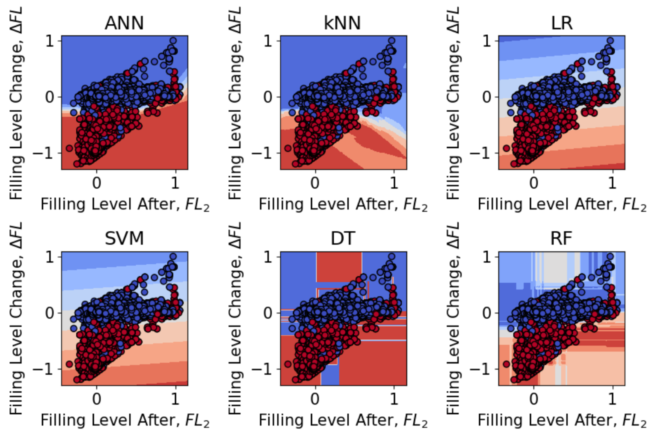


Fig. 4. Decision boundaries for the investigated classification algorithms if only filling level after (FL_2) and filling level change (ΔFL) features were used.

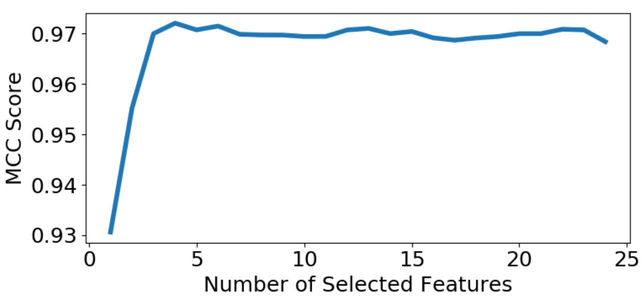


Fig. 5. The MCC score for RF using 10-fold cross-validation on the training datasets against the number of features. The best features were chosen from all the considered features using RFE.

B. Conventional classification algorithms

TABLE IV. THE PERFORMANCE COMPARISON FOR THE CONVENTIONAL CLASSIFICATION ALGORITHMS.

Algorithm	Accuracy	Recall	F1 score	MCC score
ANN	96.8 %	90.5 %	0.929	0.908
kNN	96.3 %	88.8 %	0.919	0.896
LR	96.6 %	88.3 %	0.924	0.904
SVM	96.8 %	89.3 %	0.923	0.908
DT	96.6 %	92.6 %	0.927	0.904
RF	97.2 %	91.7 %	0.938	0.920

The conventional classification algorithms were applied to the considered problem in order to see if they could improve the performance even further than the optimised manually engineered model. Six classification algorithms introduced before were investigated: ANN, kNN, LR, SVM, DT, and RF. All classification algorithms were trained and optimised using the training dataset. Similar to the manually engineered models, the dataset included three features for each time stamp: ΔFL, FL_2, and V_{str}. The obtained decision boundaries for each classification algorithm for the case of two features (FL_2 and ΔFL) are visualised in Fig. 4 . Dark red corresponds to an emptying while dark blue to a non-emptying. The fading colours reflect the decreasing likelihood that a region belongs to the corresponding class.

The trained model for each classification algorithm was evaluated on the testing dataset. The obtained performance metrics are presented in Table IV. In comparison to the optimised manually engineered model, the largest improvement in terms of the MCC score was achieved by RF (0.905 versus 0.920). Therefore, RF was chosen as the base approach for exploring the effect of including additional features to the datasets.

C. Random Forest with extended features

On the one hand, including additional features often allows achieving better classification performance. On the other hand, from the operational point of view it is preferable to use as few features as possible since it improves the interpretability of the solution. To achieve the compromise between these two conflicting requirements, *Recursive Feature Elimination* (RFE) [24] method was used to identify the best set of feature. In addition to the three original features, both training and testing datasets were extended with new features as presented in Table V.

TABLE V. EXTENDED AND SELECTED FEATURES AS THE RESULT OF RFE.

Feature	Selected
Vibration Strength	✗
Filling Level Before	✓
Filling Level After	✗
Filling Level Change	✓
Filling Level 12h Before	✗
Filling Level 12h After	✗
Filling Level Change 12h	✓
Filling Level 3h Before	✗
Filling Level 3h After	✗
Filling Level Change 3h	✓
Container Type	✗
Recyclable Fraction	✗

Note that *Recyclable Fraction* and *Container Type* are categorical features and, thus, they generated 16 binary features in total. For each size of feature set, the best features were identified calculating the MCC score for RF (number of trees was fixed to 10) using 10-fold cross-validation on the training datasets. Fig. 5 depicts the best obtained MCC scores against the number of selected features. In Fig. 5 it can be seen that the optimal number of features was four. The four chosen features are indicated in Table V.

Surprisingly, *Filling Level After* used previously was not chosen while *Filling Level Before* was among the four features. Two other new included features were variations of the *Filling Level Change* for three and twelve hours respectively.

Once the best set of features was selected, the optimal number of trees for RF was optimised on the training dataset with the selected features. The obtained results in terms of the MCC score and the training time are presented in Fig. 6. It shows that the increased number of DTs corresponds to a higher MCC score while the computational complexity grows linearly. In fact, the highest MCC score was obtained around 100 DTs in RF, however, it can be seen that there was no

TABLE VI. THE PERFORMANCE COMPARISON FOR ALL CONSIDERED SOLUTIONS.

Group	Solution	Accuracy	Recall	F1 Score	MCC score
Manually engineered model	Existing	86.8 %	47.9 %	0.633	0.608
	Optimised	96.6 %	90.5 %	0.927	0.905
ML with standard features	Artificial Neural Network	96.8 %	90.5 %	0.929	0.908
	k-Nearest Neighbours	96.3 %	88.8 %	0.919	0.896
	Logistic Regression	96.6 %	88.3 %	0.924	0.904
	Support Vector Machine	96.8 %	89.3 %	0.923	0.908
	Decision Tree	96.6 %	92.6 %	0.927	0.904
	Random Forest	97.2 %	91.7 %	0.938	0.920
ML with extended features	Random Forest	**99.1** %	**98.2** %	**0.980**	**0.974**

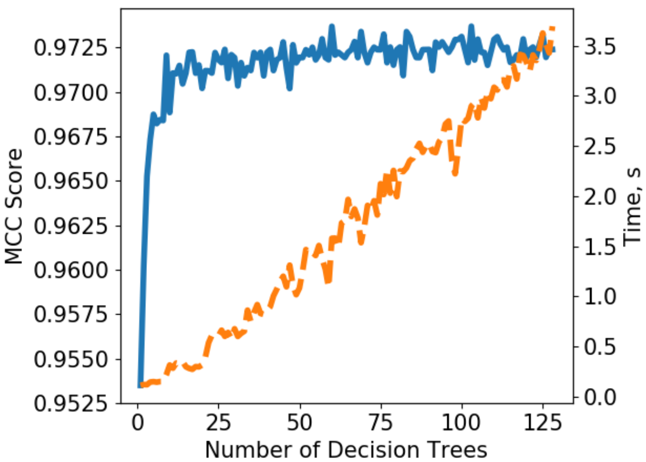


Fig. 6. Result for the hyperparameter optimisation of RF with 4 features using 10-fold cross-validation. Solid line is the MCC score. Dashed line is the training time. It can be seen that the computational complexity increases linearly while the MCC scores feature a random behaviour for more than 10 trees.

sensible difference above ten DTs. The confusion matrix on the testing dataset for the optimised RF is presented in Table VII. It classified correctly almost all time stamps in the testing dataset. Thus, the optimised RF with the extended features achieved very high classification performance: 99.1 % accuracy, 98.2 % recall, and 0.974 MCC score.

TABLE VII. CONFUSION MATRIX FOR THE OPTIMISED RANDOM FOREST TRAINED ON THE BEST SUBSET OF THE EXTENDED FEATURES.

		Actual	
		Emptying	No Emptying
Predicted	Emptying	1832	41
	No Emptying	34	6098

D. Discussion and results summary

Table VI aggregates the performance metrics for all of the considered solutions. The solutions were grouped into three categories: manually engineered model, machine learning algorithms with the same features as for manually engineered models, and the machine learning algorithm with additional features. Table VI reflects the steps in the application of the AutoML approach: 1) the evaluation of performance requires data; 2) data allows evaluating the existing solution (i.e., the existing manually engineered model); 3) data also allows optimising and evaluating the existing solution (i.e., the optimised manually engineered model); 4) the second category presents the results of conventional machine learning algorithms; 5) the third category is the result of using feature engineering.

In the first category, the existing manually engineered model performed poorly with respect to the accurate classification of the actual emptyings. This fact explains low recall, F1 and MCC scores. The observed accuracy was moderately high since the dataset is skewed towards non-emptyings. Nevertheless, a significant improvement with respect to all considered performance metrics was achieved optimising the threshold values for the manually engineered model. It signifies the importance of data analysis even as a supportive tool for the design of expert systems.

It is worth noting, that in the second category there were two types of classification algorithms: ensemble classifiers (RF) and classifiers with one model (all but RF). In terms of the fair comparison of algorithms, ensemble classifiers should be treated separately, however, in the scope of the considered problem the achieved performance was the main goal. The use of single classifiers did not provide qualitative performance improvement in comparison to the optimised manually engineered model. In fact, it was only ANN where all performance metrics were marginally greater than or equal to the optimised manually engineered model. Other single classifiers were either better in some metrics but worse in others (SVM and DT) or worse in all the metrics (kNN and LR). Expectedly, the largest gain was observed for the ensemble classifier (RF). Thus, in the second category the only improvement was achieved with the use of the ensemble-based algorithm.

Extended features in the third category combined with the ensemble-based algorithm allowed moving the performance even further. In comparison to the existing manually engineered model, this best performing solution improved the accuracy by 12.3 percentage points, the recall by 50.3 percentage points, the F1 score by 0.347, and the MCC score by 0.366. Overall, we see that applying each step of the methodology allowed improving the performance of solutions.

Recall that the dataset includes time stamps for different container types. It was observed that there were no container types with low classification performance and the deviations were not large. Thus, the best performing solution worked equally well for different container types. It also explains why the *Container Type* feature was not chosen by RFE.

Finally, in order to evaluate the effect of the improved emptying detection on quality of filling level predictions, we

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

performed an experiment on the testing dataset using the detections of the existing manually engineered model and the ones obtained from the best performing solution. During the experiment, we fitted the regression model (cf. Fig. 1) to the filling levels measurements in order to predict time when the filling level of a recycling container would reach 90.0 % (i.e., when it should be emptied). The forecasted time is considered successful if the filling level would reach 90.0 % either 72 hours before or 24 hours after the prediction[2] otherwise it is considered that the prediction failed. The performance of a model is defined as the ratio of the successful forecasts to the total number of forecasts in the testing dataset. The performance of the regression model obtained on the emptying detection provided by the best performing solution was 14.2 percent higher than the result achieved for the existing manually engineered model. This is an important quantitative characterisation demonstrating the significance of the correct detection of emptyings for the whole system.

VI. Conclusion

This article presented the use of automated machine learning approach for industrial informatics with the show-case of the accurate detection of emptying a recycling container using the measurements from the sensor mounted on top of the container. The article proposed the iterative data-driven methodology for achieving the highest performance where first the existing solution to the problem was assessed, second this solution was optimised using the collected dataset, next, machine learning algorithms were applied to the problem, and finally, the feature engineering was used to find if additional features would improve the results. There are several limitations to this study. First, it was not directly quantifying to which extend inaccurate detection of emptying affects filling level predictions. Second, the investigated solutions assume availability of filling level measurements and vibration strength scores. Third, the computational complexity of the investigated solutions was not taken into account. During the investigation, several solutions were considered: the existing manually engineered model, the optimised manually engineered model, conventional machines learning algorithms and conventional machine learning algorithms with the extended features. On the obtained dataset, the existing manually engineered model demonstrated a moderate accuracy of 86.8 % but a poor recall of 47.9 %. However, the recall of the manually engineered model was improved to 90.5 % by simply optimising the model's parameters using the grid search on the training dataset. Notably, the combination of the extended features and the ensemble classifier (Random Forest) allowed significantly improving the performance reaching the accuracy of 99.1 % and the recall of 98.2 %. The set of extended features was formed using Recursive Feature Elimination algorithm and included the filling level before a potential emptying as well as three different filling level changes: the immediate change, the change in three hours, and the change in twelve hours. Moreover, we performed the experiment assessing the effect of improved emptying detection on the emptying time prediction. The results demonstrated that the detections of the best performing solution improved the emptying time predictions by 14.2 percent compared to the detections of the existing manually engineered model. It is concluded that the best performing solution is applicable for a practical deployment in the production system if the requirement on the availability of the measurements preceding and succeeding the point of interest by twelve hours is relaxed. The achieved results emphasise the importance of the data-driven engineering for the design of Smart Waste Management systems.

[2]The limits are justified by the planning of the truck's routes. It usually happens either a day before or in the morning. Thus, currently, it is sufficient to be able to forecast within a time window. The prediction limit is not symmetric because it is safer to provide a prediction prematurely (overfilling is not desirable).

References

[1] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter, "Efficient and Robust Automated Machine Learning," in *Advances in Neural Information Processing Systems 28*, 2015, pp. 2962–2970.

[2] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2013, pp. 847–855.

[3] European Commission, "Waste," http://ec.europa.eu/environment/waste/, [Online; accessed 23-May-2018].

[4] T. Anagnostopoulos, A. Zaslavsky, K. Kolomvatsos, A. Medvedev, P. Amirian, J. Morley, and S. Hadjieftymiades, "Challenges and Opportunities of Waste Management in IoT-Enabled Smart Cities: A Survey," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 3, pp. 275–289, 2017.

[5] "Bigbelly - Smart City Solutions," http://bigbelly.com/, [Online; accessed 27-May-2018]. [Online]. Available: http://bigbelly.com/

[6] "Ecube Labs - Smart waste management solution," https://www.ecubelabs.com/, [Online; accessed 27-May-2018]. [Online]. Available: https://www.ecubelabs.com/

[7] "Enevo - Waste and Recycling Services Right-Sized for You," https://enevo.com/, [Online; accessed 27-May-2018]. [Online]. Available: https://enevo.com/

[8] "Sensoneo - Smart Waste Management solution," [Online; accessed 27-May-2018]. [Online]. Available: https://sensoneo.com

[9] "Onlab - Onsense," http://www.onlab.com.tr/en/onsense_eng/, [Online; accessed 27-May-2018]. [Online]. Available: http://www.onlab.com.tr/en/onsense_eng/

[10] "Smart Waste — Citibrain: Smart Integrated Solutions for Smart Cities," http://www.citibrain.com/en/solutions/smart-waste/, [Online; accessed 27-May-2018]. [Online]. Available: http://www.citibrain.com/en/solutions/smart-waste/

[11] "SmartBin IoT Level Sensors — Intelligent waste monitoring," https://www.smartbin.com/solutions/iot-level-sensors/, [Online; accessed 27-May-2018]. [Online]. Available: https://www.smartbin.com/solutions/iot-level-sensors/

[12] B. Zhang, C. H. Liu, J. Tang, Z. Xu, J. Ma, and W. Wang, "Learning-Based Energy-Efficient Data Collection by Unmanned Vehicles in Smart Cities," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1666–1676, 2018.

[13] B. Tang, Z. Chen, G. Hefferman, S. Pei, T. Wei, H. He, and Q. Yang, "Incorporating Intelligence in Fog Computing for Big Data Analysis in Smart Cities," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2140–2150, 2017.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2915572, IEEE Transactions on Industrial Informatics

[14] V. Catania and D. Ventura, "An Approch for Monitoring and Smart Planning of Urban Solid Waste Management Using smart-M3 Platform," in *Proceedings of the 15th Conference of Open Innovations Association FRUCT*, ser. FRUCT'15. Helsinki, Finland, Finland: FRUCT Oy, 2014, pp. 4:24–4:31. [Online]. Available: https://doi.org/10.1109/FRUCT.2014.6872422

[15] S. Sharmin and S. T. Al-Amin, "A cloud-based dynamic waste management system for smart cities," in *Proceedings of the 7th Annual Symposium on Computing for Development*, ser. ACM DEV '16. New York, NY, USA: ACM, 2016, pp. 20:1–20:4. [Online]. Available: http://doi.acm.org/10.1145/3001913.3006629

[16] J. W. Lu, N. B. Chang, L. Liao, and M. Y. Liao, "Smart and green urban solid waste collection systems: Advances, challenges, and perspectives," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2804–2817, Dec 2017.

[17] T. S. Vasagade, S. S. Tamboli, and A. D. Shinde, "Dynamic solid waste collection and management system based on sensors, elevator and GSM," in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, March 2017, pp. 263–267.

[18] M. H. A. Wahab, A. A. Kadir, M. R. Tomari, and M. H. Jabbar, "Smart Recycle Bin: A Conceptual Approach of Smart Waste Management with Integrated Web Based System," in *2014 International Conference on IT Convergence and Security (ICITCS)*, Oct 2014, pp. 1–4.

[19] B. Chowdhury and M. U. Chowdhury, "RFID-based real-time Smart Waste Management system," in *2007 Australasian Telecommunication Networks and Applications Conference*, Dec 2007, pp. 175–180.

[20] J. Lindström, A. Hermanson, M. Hellis, and P. Kyösti, "Optimizing recycling management using industrial internet supporting circular economy: A case study of an emerging IPS^2," *Procedia CIRP*, vol. 64, pp. 55 – 60, 2017, 9th CIRP IPSS Conference: Circular Perspectives on PSS. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2212827117302184

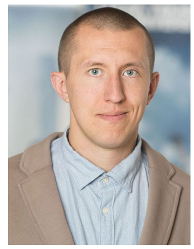
[21] J. Lindström, A. Hermanson, F. Blomstedt, and P. Kyösti, "A multi-usable cloud service platform: A case study on improved development pace and efficiency," *Applied Sciences*, vol. 8, 2018. [Online]. Available: http://www.mdpi.com/2076-3417/8/2/316

[22] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson Education, Inc., 2010.

[23] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.

[24] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.

David Rutqvist received the M.S. degree in computer science from the Luleå University of Technology, Luleå, Sweden, in 2018. He has been a technical architect and software developer at BnearIT AB, Luleå, Sweden from 2013 to 2018. Since 2018 he is a Consultant at Netlight Consulting AB, Stockholm, Sweden. His research interests include software architecture, distributed systems, and machine learning, especially in the application to the Internet of Things.



Denis Kleyko received the B.S. degree (Hons.) in telecommunication systems and the M.S. degree (Hons.) in information systems from the Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia, in 2011 and 2013, respectively, and the Ph.D. degree in computer science from the Luleå University of Technology, Luleå, Sweden, in 2018, where he is currently a Postdoctoral Researcher with the Dependable Communication and Computation Systems Group, Department of Computer Science, Electrical and Space Engineering. His current research interests include vector symbolic architectures, hyperdimensional computing, artificial intelligence, cognitive architectures, and machine learning.

Fredrik Blomstedt received the M.S. degree in computer science from the Luleå University of Technology, Luleå, Sweden, in 1997. He is a technical architect and owner at BnearIT AB, Luleå, Sweden since 2002. His research interests include software architectures, Industrial Internet and Internet of Things as well as development processes and operational excellence.