Application of machine learning algorithms in municipal solid waste management: A mini review

Wanjun Xia1,2, Yanping Jiang2, Xiaohong Chen2 and Rui Zhao3

Abstract
Population growth and the acceleration of urbanization have led to a sharp increase in municipal solid waste production, and researchers have sought to use advanced technology to solve this problem. Machine learning (ML) algorithms are good at modeling complex nonlinear processes and have been gradually adopted to promote municipal solid waste management (MSWM) and help the sustainable development of the environment in the past few years. In this study, more than 200 publications published over the last two decades (2000–2020) were reviewed and analyzed. This paper summarizes the application of ML algorithms in the whole process of MSWM, from waste generation to collection and transportation, to final disposal. Through this comprehensive review, the gaps and future directions of ML application in MSWM are discussed, providing theoretical and practical guidance for follow-up related research.

Keywords
Municipal solid waste management, machine learning, deep learning, sustainable development, data-driven

Received 16th March 2021, accepted 1st July 2021 by Associate Editor Nemanja Stanisavljevic.

Introduction
Municipal solid waste (MSW) indicates the solid and semi-solid waste generated in urban daily life, including residential, industrial, institutional, commercial, municipal, and construction and demolition waste (Ofori-Boateng et al., 2013; Yi et al., 2011). With rapid urbanization, economic development, and population growth, MSW has significantly increased (Magazzino et al., 2020; Mukherjee et al., 2020), and the amount of MSW worldwide will increase to 2.2 billion tons per year by 2025 (Hoornweg and Bhada-Tata, 2012). The large amount of MSW has become a serious threat to the city and its surrounding ecological environment, which causes many issues, such as illegal dumping and pollution (Da Paz et al., 2020; Demirbilek et al., 2013; Magazzino et al., 2020; Nagpure, 2019; Santos et al., 2019). MSW is considered a major global environmental problem, especially in developing countries (Haraguchi et al., 2019; Nguyen et al., 2020). Therefore, developing efficient MSW management (MSWM) is crucial for protecting resources, the environment, and public health (Ceylan et al., 2020), but the environmental issues of MSW are normally difficult to solve because of their diversity and heterogeneous nature (Bagheri et al., 2019).

With this background, waste management services have focused on the use of advanced information technology to improve the MSWM and promote the efficiency of waste sorting and recycling in recent years (Nowakowski and Pamula, 2020; Rahman et al., 2020). Owing to their excellent ability to model complex mechanisms, machine learning (ML) methods have been successfully applied to environmentally related fields such as wastewater, air pollutants, and solid waste treatments (Joharestani et al., 2019; Ye et al., 2020). The MSWM system includes a variety of nonlinear relationships, and the traditional linear model is limited (Chhay et al., 2018). Therefore, ML methods are expected to play an important role in the modeling, prediction, and optimization of MSW-related issues. In previous studies, Ye et al. (2020) discussed the various artificial intelligence (AI) models used in solid waste applications based on 85 studies published between 2004 and 2019, and Li et al. (2020) introduced information technologies used in construction and demolition waste management applications. In contrast to these works, this paper mainly focuses on the application of the ML algorithm in MSWM and expounds upon it from the perspective of the whole process of waste management.

1School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu, Sichuan, China
2Library, Southwest Jiaotong University, Chengdu, Sichuan, China
3Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, Sichuan, China

Corresponding author:
Wanjun Xia, School of Computing and Artificial Intelligence, Southwest Jiaotong University, West Park of Hi-Tech Zone, Chengdu, Sichuan 611756, China.
Email: xiawanjun@swjtu.cn
Overview of ML

ML is a multidisciplinary field that covers computer science, probability theory, statistics, approximate theory, and complex algorithms, and its theories and methods have been widely used to solve complex problems in engineering applications (Sanchez-Lengeling and Aspuru-Guzik, 2018; Zhang et al., 2020). ML uses data to “train” and learns how to complete tasks from the data through various algorithms. These algorithms attempt to mine hidden information from a large amount of historical data and use them for regression or classification. ML algorithms include artificial neural networks (ANNs), support vector machine (SVM), naive Bayes, K-nearest neighbor (KNN), decision tree (DT), random forest (RF), and adaptive network fuzzy inference system (ANFIS). As the most important branch of ML, deep learning (DL) has developed rapidly in recent years and has gradually become a research hotspot in ML. The workflow of the traditional ML methods is shown in Figure 1, and it usually consists of three steps: (1) data processing and feature extraction, (2) choosing the proper ML algorithms and parameters, and (3) testing and evaluating performance. Table 1 lists the most frequently used ML models in the MSWM.

Data and methods

Literature related to the MSWM were obtained from Web of Science, specifically from the sub-databases of Science Citation Index Expanded, Social Sciences Citation Index, and Arts & Humanities Citation Index. Considering that many of the leading ML models were only published as conference papers, the CPCI-S database was also included in the search scope.

The search strategy was: Topic = municipal waste and Topic = “machine learning” or “deep learning” or “neural network*” or “support vector machine” or SVM or “support vector regression” or SVR or ANFIS or “logic regression” or “naive Bayes*” or “decision tree” or “k means” or KNN or “k nearest neighbor*” or “recurrent neural network” or RNN or “convolutional neural network” or CNN or “long short term memory” or LSTM or “random forest*” or “gradient boosting decision tree” or GBDT or “gradient boosting regression tree” or GBRT. The retrieval time was limited to 2000–2020. After a preliminary search, 229 studies were obtained and screened manually. Articles were required to meet the following criteria: (1) involved the application of ML techniques for MSWM, (2) had a clear description of the ML model used, and (3) language was limited to English. Review articles, wastewater treatment, and others unrelated to MSWM were not included in this paper. A total of 137 articles were selected. Meanwhile, we expanded the literature data through references and citing papers, and finally, 226 related articles were selected. Figure 2 shows the distribution of related studies over the publication year, and it can be seen that relevant publications in this field are increasing gradually, especially in the last 2 years. Among them, DL methods have been used with increasing frequency. Data-driven methods have received increasing attention from researchers, and the traditional ML and DL provide an effective way to solve the practical problems of MSWM.

In general, ML algorithms have been widely used in MSWM, from waste generation to collection and transportation. They are also integrated into various waste disposal processes such as landfill, composting and incineration to assist energy recovery and gas treatment, which can help achieve sustainable development through low-risk waste reduction and resource utilization. Figure 3 shows the application of the ML algorithms in each stage of the MSWM.

Application of ML algorithms in MSWM

Waste generation prediction

The equipment and facilities related to the disposal of MSW often require high investment, and authorities should not only focus on the current situation but also plan and deploy in advance based on the MSW generation trend. Reliable prediction of MSW generation is very important because it can provide data support for urban environmental planning, operation, and supervision and can aid in developing reasonable implementation plans for waste collection, transportation, and treatment (Oguz-Ekim, 2021). Therefore, MSW forecasting is the basis for waste management. However, MSW generation is complicated, and it is
affected by many factors such as demography, economic growth, and individual behaviors (Ceylan, 2020). The relationship between the influencing factors and MSW generation is not simply linear, and nonlinear models have shown better accuracy than linear models (Noori et al., 2009; Wu et al., 2020a). Many ML methods have been developed for MSW generation forecasting, with satisfactory results. Considering the different algorithms used and their application in different countries or regions, Table 2 provides a summary of the ML methods used to predict MSW generation. According to the length of the forecast period, current studies usually can be divided into short-term prediction (from weekly to monthly), mid-term prediction (months to 3–5 years), and long-term prediction (many years ahead) (Ye et al., 2020). The prediction scale ranges from homes, buildings to cities, and countries.

According to the articles reviewed, different ML algorithms such as ANN, SVM, and ANFIS, have been used to build the prediction model, and factors such as weather data and demographic and socioeconomic data are often chosen as input features. Among them, ANN is the most commonly used method, but its performance is affected by data over-fitting, local minima, and poor generalization (Abbasi and El Hanandeh, 2016). However, GBRT is composed of multiple regression trees with strong generalization and is easy to interpret, and often achieves better results than ANN (Kontokosta et al., 2018). Moreover, MSW generation can be analyzed by time series, and currently,
DL algorithms, such as LSTM, have performed well in sequence prediction (Chang et al., 2020). Therefore, several studies have used LSTM to forecast MSW generation with high accuracy, especially when using large amounts of data (Cubillos, 2020; Huang et al., 2020b).

Although ML models can accurately predict MSW generation, the biggest challenge is the lack of historical data and other necessary data, especially at the level of subdivisions such as households, buildings, or communities (Cubillos, 2020), which can be attributed to ineffective waste management (Ayeleru et al., 2021; Solano Meza et al., 2019). Given sufficient data, ML methods may achieve a higher accuracy in MSW generation prediction (Kannangara et al., 2018).

### Waste collection and transportation

The MSW collection and transportation system, as a hub connecting the source, final disposal, and resource recovery management, is the key to ensuring the normal flow of urban material and energy. On the one hand, the rapid increase in the generation of MSW has led to the failure of back-end treatment. On the other hand, the operation scope of MSW collection and transportation systems increases rapidly, and the complexity of the management also increases accordingly. Therefore, optimizing the collection and transportation route of MSW can effectively reduce transportation costs and improve operational efficiency.

#### Waste bin detection

Waste collection faces problems of high transportation costs, and trucks often visit waste bins that are only partially full, which is an inefficient use of resources (Pereira Ramos et al., 2018). Many factors can affect waste accumulation, so it is difficult to predict the fill levels of waste bins. To improve the efficiency of waste collection, several studies have applied ML algorithms to realize waste bin detection, which can be regarded as a classification problem. The input data often come from the level of image sensors installed in a waste bin, and feature extraction is carried out based on the image processing technique. Hannan et al. (2014) used an ANN with a Hough transform model or Gabor wavelet filters to classify the level of solid waste inside the bin, and the output was divided into five classes: empty, medium, full, flow, and overflow. Combined with gray-level co-occurrence matrix feature extraction methods, Arebey et al. (2012) adopted ANN and KNN for solid waste bin level detection and found that the KNN classifier performed better than ANN with the same dataset. Based on the data collected from the sensor mounted inside the container, Rutqvist et al. (2020) treated the waste bin level detection as a two-class classification problem: emptying and non-emptying, and ANN, KNN, LR, SVM, DT, and RF were selected as classification algorithms, among which RF performed best (F1 score reached 0.938). Compared to the existing manually engineered model, ML methods improve the classification accuracy, and intelligent detection of waste bin levels can reduce the driving distance of trucks, which could promote money savings for the management agencies, and reduce carbon dioxide emissions (Vecchi et al., 2016).

#### Routing optimization

The cost of waste collection and transportation accounts for 60%-80% of the total waste management system costs (Wu et al., 2020b), so optimizing the route of vehicles for waste collection and transportation can save time, reduce the running distance, vehicle maintenance cost, and fuel cost, and can also effectively arrange vehicles and allocate human resources. A few studies have applied ML algorithms to waste collection routes. Hoang et al. (2019a) combined a time-series ANN nonlinear autoregressive model with geographic information system (GIS) to analyze the effects of changes in waste composition and density on truck route optimization. For example, Ferreira et al. (2017) developed an ANN model to predict the collection frequency for each location, which can reduce unnecessary visits. In addition, few other studies combined a clustering approach in vehicle route optimization of waste collection (Al-Refaie et al., 2020; Jammeli et al., 2019; Wu et al., 2020b), namely “cluster first, route second.” First, the K-means algorithm was used to aggregate the bin locations into clusters and determine the minimum number of collection vehicles or the
| Table 2. Summary of ML algorithms used to predict MSW generation. |
|---------------------------------------------------------------|
| **Input feature** | **Output** | **Performance evaluation** | **Scale** | **Advantage** | **Disadvantage** | **Reference** |
| Weekly household | Weekly household | RMSE: 0.034 (GBRT) | City | Low accuracy | Easy to overfitting | Ayeleru et al. (2021) |
| Waste generation rate in New York, USA | Weekly municipal waste generation in New York, USA | Weekly plastic waste generation rate in Dhanbad, India | Household | Use multi-site models to increase amounts of historical data | Use feature selection techniques | Kontokosta et al. (2018) |
| Education, occupation, income, house type | Monthly and Seasonal municipal waste generation | Mid-term ANN 171, 130, and 48 for each region | Region | Evaluate the effect of each predictor in detail | No clear comparison with other papers | Antanasijevic et al. (2013) |
| GDP per Capita, domestic material consumption, resource productivity | GDP per Capita, domestic material consumption, resource productivity | Monthly and Seasonal municipal waste generation | Country | High accuracy | Low accuracy | Antanasijevic et al. (2013) |
minimum distance traveled by collection vehicles. In the second step, an optimization algorithm is used to solve the routing problem for each cluster. The result of the optimization is the reduction of travel distance and pollution emissions, which is good for the economy and environment.

Waste classification. Waste classification is a scientific management method for effective waste disposal. Through classification, useful materials are recycled and utilized, and the quality of waste is improved, which is convenient for landfill or incineration to promote treatment that yields fewer risks to human health and the environment. Moreover, the classification and reuse of recyclable MSW is important for a circular economy. Traditional waste classification mainly relies on manual selection (Huang et al., 2020a), which is inefficient, and with the development of artificial intelligence, many ML algorithms have been proposed to improve the accuracy of recyclable MSW identification (Liu et al., 2020; Ping et al., 2020). In recent years, CNNs have achieved remarkable results in image classification. Therefore, by taking pictures of MSW, the CNN can automatically identify different types of waste. Table 3 summarizes several CNN models used for the intelligent classification.

At present, various CNN architectures have been used (e.g. VGGNet, AlexNet, ResNet, and DenseNet) to build the MSW classification model, which has high accuracy. However, one of the major challenges is that there are few waste image datasets available for model training, and Yang and Thung (2016) established the public dataset TrashNet, which contained six categories and a total of 2527 images. However, there is still a lack of large-scale databases such as ImageNet. A small training dataset cannot accurately capture the characteristics of various types of MSW. To overcome the drawback of insufficient data, data augmentation methods, such as horizontal or vertical flipping, have been adopted to increase the image amount. Another challenge is that CNN has a large number of parameters, and the training process requires considerable resources and time; therefore, transfer learning has been used to share the trained model parameters with the new one to improve and optimize the learning efficiency, which is the state-of-the-art algorithm for waste classification (Figure 4). In particular, Mao et al. (2020) proposed optimized DenseNet121, obtaining an accuracy of 99.6% on the TrashNet dataset, which is the best result for this public dataset until now, but the training time was slightly slow at 5542 seconds. Vo et al. (2019) designed a lightweight model, RecycleNet, that yielded prediction times of only 15.9 ms (GPU) and 352 ms (CPU). However, the accuracy of RecycleNet was relatively low with only 81%. Therefore, future research should improve the effectiveness of the proposed framework in actual systems and achieve a balance between running time and accuracy.

Waste treatment and disposal

Composting. Composting is a valuable method for organic waste treatment (Walling et al., 2020). Organic matter is the main fraction of MSW; therefore, composting can be considered a cost-effective option for MSW disposal. After composting, MSW becomes a hygienic and odorless humus, realizing the key aspects of harmlessness, waste reduction, and recycling. ML algorithms can help model the complex processes that occur during composting, such as maturity prediction, parameters control, and optimization (Table 4).

Maturity prediction is often regarded as a classification task. Input features include a series of physical, chemical, and biological indicators, such as color, texture, temperature, water content, carbon-to-nitrogen ratio (C/N), and humus content (Kujawa et al., 2014). Few studies have attempted to use traditional ANN and CNN to learn the morphology, texture, and color characteristics of compost images and establish compost maturity prediction models (Kujawa et al., 2020; Xue et al., 2019). The composting process is complex and affected by many factors, such as temperature, pH, aeration, and C/N. During waste composting, some toxic and harmful gases such as CH₄ and N₂O are released, which can cause environmental pollution and may adversely affect human health. To control and optimize the composting process, some researchers have proposed ML methods to predict the response parameters and model gas emissions (Díaz et al., 2012). Generally, the application of ML in composting is in the exploratory stage, and almost all the algorithms currently used are ANN, and the data are difficult to obtain, mainly from experimental tests, so some development is necessary before practical application.

Incineration. Incineration has the advantages of requiring a small land area, short treatment time, and good energy reclamation. It can reduce waste volume by 90% and mass by 70%–80%, which has become the main method of MSW disposal. Owing to the heterogeneity of MSW, its energy content can be recovered through thermochemical processes (Drudi et al., 2019); therefore, incineration is also an effective method of energy generation, realizing the conversion from waste to energy. However, the use of incineration can produce harmful gas pollutants, including dust, toxic heavy metals, acidic gases (e.g. NOₓ and SOₓ), and organic pollutants (e.g. dioxins). Therefore, flue gas treatment has a significant impact on the implementation of waste incineration.

With the development of information technology, ML methods have been used in incineration, such as heating value (HV) prediction and pollutant emission monitoring (Table 5). The ANN is still the most commonly used model for HV prediction. The input features are often based on ultimate analysis, physical composition, or proximate analysis, and higher heating value (HHV) or lower heating value (LHV) is the output of the prediction model. Although these models claimed to have high accuracy, they used only one or two hundred data points, which were prone to overfitting. Moreover, these methods inherited the shortcomings of the experimental method, and the input features were derived from the ultimate analysis or approximate analysis, which was considered inaccurate because of the huge difference.
| Model | Classification output | Data size | Accuracy (%) | Advantage | Disadvantage | Reference |
|-------|------------------------|-----------|--------------|-----------|-------------|-----------|
| CNN [VGGNet, DenseNet and NASNetLarge] | 6-classification: paper, glass, metal, plastic, textile, and organic waste | TrashNet [2527 images] and manually collect (5000 images) | 96.5 and 94 | Combine different candidate classifiers to improve accuracy | Limit amount of data | Huang et al. (2020a) |
| CNN [Improved DenseNet] | 6-classification: cardboard, glass, metal, paper, plastic, trash | TrashNet + data augmentation [10,108 images] | 99.6 | Optimize fully connected layer of CNN | Lack of comparison experiments | Mao et al. (2020) |
| CNN [Improved ResNet] | 6-classification: paper, glass, metal, plastic, textile, and organic waste 3-classification: organic, inorganic, medical waste | TrashNet [2527 images] and internet resources [5904 image] | 94 and 98 | Compare the classification accuracy of multiple models | Limit amount of data | Vo et al. (2019) |
| CNN [Improved DenseNet] | 6-classification: cardboard, glass, metal, paper, plastic, trash | TrashNet [2527 images] + data augmentation | 81 | High efficiency | The accuracy is not high | Bircanoğlu et al. (2018) |
| AutoEncoder + CNN [AlexNet, GoogLeNet and ResNet]+ Ridge Regression +SVM | Organic or recyclable waste | Open-access dataset [25,077 waste images] | 99.95 | High accuracy | High complexity | Togacar et al. (2020) |
| CNN [AlexNet]+MLP | Recyclable or the others | Manually collect [5000 images] | 98.2 [fixed orientation] 91.6 [random orientations] | Higher accuracy than using CNN alone | Lack of comparison experiments | Chu et al. (2018) |

Accuracy refers the ratio of the number of correctly classified samples to the total number of samples.
in the composition of MSW. To overcome the shortcomings of experimental methods, You et al. (2017) analyzed the operating conditions of incinerators to obtain the variation trend of the HV. ANN, SVM, ANFIS, and RF were used as training models; among them, ANFIS performed best, although it consumed a longer time. Birgen et al. (2021) developed a Gaussian process regression (GPR) model to predict the daily LHV of MSW using historical data from a waste-to-energy plant, together with weather and calendar data, to realize the online prediction of LHV. Similarly, ML models can be used to realize the online measurement of dioxin and serve as a good tool to monitor waste incineration equipment under real-time conditions (Behera et al., 2014; Shaha et al., 2020). Intelligent control and prediction can maximize energy production and minimize environmental impact, achieving more stable combustion and cleaner flue gas emissions.

**Landfill.** Landfills are widely used waste disposal methods at home and abroad. Waste that cannot be recycled or incinerated is placed in landfills (Edgar et al., 2020). This method can handle large amounts of MSW, which is convenient and easy to operate. However, landfills occupy a large amount of land resources, and in the process of waste decomposition, various physical, chemical, and biological reactions occur in the landfill, resulting in serious pollution. As shown in Table 6, some studies have used ML algorithms to estimate landfill areas, predict landfill leachate generation, and monitor landfill gas (LFG). ANN and ANFIS are the most commonly used models. Leachate and LFG generation are often affected by meteorological factors and landfill design. ML algorithms use influencing factors such as temperature and rainfall as input features to model the process of leachate and LFG production and obtain high accuracy (Abunama et al., 2019; Li et al., 2011). An accurate and reliable prediction is needed for proper design and operation and helps eliminate environmental impacts and maximize the use of resources.

**Discussion**

Although ML algorithms have been widely used in MSWM, they are still at an early stage of development and use. There are some gaps in the application of ML algorithms to MSWM, and some aspects should be considered in the future.

**Data collection**

ML methods search for hidden information from a large amount of data using different algorithms. Sufficient and reliable data were the most basic and core parts. The historical length of the waste data and its quality are critical for reliable ML model performance (Chen and Lin, 2008; Masebinu et al., 2017). However, most current MSWM research deals with small datasets (Díaz et al., 2012; Hosseinizadeh et al., 2020; Ozkan et al., 2015), and this may be attributed to waste management infrastructure and practices (Ayeleru et al., 2021). Waste-related data are managed by different channels involving several stakeholders, making data collection and compilation difficult (Abbasi and El Hanandeh, 2016). Due to insufficient data, accurate ML models related to MSWM are difficult to build (Abunama et al., 2019; Bagheri et al., 2019). Furthermore, nearly all datasets used for training are not publicly available and, without an open benchmark dataset, will hinder the application of ML algorithms in MSWM.

At present, there are several ways to solve the problem of small data. For example, in waste classification, data augmentation techniques are used to increase the sample data to better apply the DL algorithm (Bircanoglu et al., 2018; Mao et al., 2020).
Table 4. Summary of ML algorithms used in MSW composting.

| Application                              | Algorithms       | Data size | Input feature                                                                                      | Output                                                                 | Performance evaluation                     | Reference            |
|------------------------------------------|------------------|-----------|---------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|--------------------------------------------|----------------------|
| Maturity prediction                      | Wavelet Neural Network | 500       | High temperature duration, moisture content, volatile solids, the value of fecal bacteria and germination index (5 input parameters) | 4-classification: full maturity, preferable maturity, general maturity, immaturity | MSE: 0.066                                | Gao et al. (2007)    |
|                                          | ANN              | 1536      | Color and texture features [input parameters range from 14 to 49]                                 | 2-classification: early maturity or not                                 | Classification error: 1.56%                | Kujawa et al. (2014) |
|                                          | CNN              | 1312      | RGB images                                                                                       | 2-classification: probability of maturity classes                       | Classification error: 0.51% to 17.77%      | Kujawa et al. (2020) |
| Composting process control and optimization | ANN              | 412       | Time, flow direction, density indicator, hydraulic load, flow [5 input parameters]              | Pressure drop                                                          | Correlation coefficient: 0.906             | Sidelko et al. (2019) |
|                                          | ANN              | 550       | Chemical and physical parameters of composting [7 input parameters]                             | Ammonia emission                                                       | R: 0.972–0.981                            | Boniecki et al. (2012) |
|                                          | ANN              | 52        | Food and yard percentage, ash and scoria percentage, the moisture content, the fixed carbon content, the total amount of organic matter, high calorific value, and pH [7 input parameters] | C/N                                                                    | Average relative error: 6.376%            | Bayram et al. (2011) |
|                                          | ANN              |           | Time, turning frequency and mixing ratio [3 input parameters]                                   | Response parameters: temperature, pH, O$_2$, respiration index, total organic carbon, TN, TP | RSME: range from 4.1 to 8.3%              | Soto-Paz et al. (2020) |
|                                          | ANN              | 20        | pH, electrical conductivity, C/N, NH$_4$/NO$_3$ ratio, water soluble carbon, dehydrogenase enzyme, TP or TN [7 input parameters] | R$^2$ (TN): 0.9983 R$^2$ (TP): 0.9991                                    |                                            | Hosseinzadeh et al. (2020) |
| Application                  | Algorithms         | Data size | Input feature                                                                 | Output | Performance evaluation                                                                 | Reference                |
|-----------------------------|--------------------|-----------|--------------------------------------------------------------------------------|--------|----------------------------------------------------------------------------------------|--------------------------|
| HV prediction               | GPR                | 1024      | Temperature, precipitation, wind strength, day of the week, week of the year [5 input parameters] | LHV    | MAPE: 5.23%                                                                            | Birgen et al. (2021)     |
|                             | ANN, SVM, ANFIS, and RF | 2200      | Feeding rate of MSW and coal, bed temperature, furnace outlet gas temperature, etc. [11 input parameters] | 9-class of LHV | Prediction precision: ANFIS [96%] > RF [90%] > SVM [87.5%] > ANN [73.5%]            | You et al. (2017)        |
|                             | ANN                | 89        | Food, paper, plastic, textile, wood, moisture [6 input parameters]             | LHV    | R: 0.993                                                                               | Ozveren (2016)           |
|                             | SVM, ANN           | 252       | C, H, O, S, H [5 input parameters]                                            | HHV    | [SVM] R²: 0.973 RMSE: 1.44 [ANN] R²: 0.978 RMSE: 1.25                                | Bagheri et al. (2019)    |
| Pollutants emission         | ANN                | 123       | C, O, H, N, S, ash and moisture content [7 input parameters]                  | HHV    | R: 0.986 RMSE: 3.095                                                                   | Olatunji et al. (2019)   |
| monitoring                  | ANFIS              | 1000      | Temperature of first and second hearth, Temperature at outlet secondary chamber, Concentration of O₂ in flue gas [3 input parameters] | Carbon monoxide emission | R²: 0.98 RMSE 4.45                                                                    | Norhayati and Rashid, (2018) |
|                             | ANN                | –         | Temperature, CO, HCl, PM [range from 3 to 5 according to different type of incinerators] | PCDDs + PCDFs | MSE [for different incinerator]: 37.43, 45.98, 0.1667 | Chang and Chen (2000)    |
| RPF-fired boiler monitoring | RF, LSTM, ANN      | 215       | Conveyor speed, feed water temperature and pressure, and incinerator exit temperature [4 input parameters] | Steam temperature, steam pressure, and steam flow rate | MAPE: RF [1.12, 5.54, 15.6] LSTM [1.14, 4.21, 9.63] ANN [6.93, 1.49, 18.77] | Behera et al. (2014); Shaha et al. (2020) |
Table 6. Summary of ML algorithms used in MSW landfill.

| Application                  | Algorithms       | Data size | Input feature                                                                 | output                           | Error metrics               | Reference                        |
|------------------------------|------------------|-----------|-------------------------------------------------------------------------------|----------------------------------|------------------------------|----------------------------------|
| Landfill leachate prediction | ANN, SVM         | 120       | Waste quantity, rainfall level, and emanated gases, etc. (9 or 3 input parameters) | Monthly leachate generation rate | $R^2$: 0.964                | Abunama et al. (2019)           |
|                              | ANFIS            | 120       | Waste quantity, rainfall level, and emanated gases [3 input parameters]        | Monthly leachate generation rate | $R$: 0.952                  | Abunama et al. (2018)           |
|                              | ANN              | –         | Meteorological parameters, measured parameters [11 input parameters]           | Leachate daily flow-rate         | $R$: 0.847                  | Karaca and Ozkaya (2006)        |
| Landfill gas prediction      | Auto-regressive  | 1883      | Mean air temperature, maximum pressure, minimum humidity, maximum wind speed, maximum humidity [5 input parameters] | Daily CH$_4$ generation rate     | MAPE 3.03%                  | Fallah et al. (2020)           |
|                              | neural network   |           |                                                                               |                                  |                             |                                  |
|                              | ANN              | 121       | Soil and air temperature, soil moisture content, CH$_4$ and O$_2$ concentration [5 input parameters] | CH$_4$ oxidation                 | $R^2$: 0.937                | Abushammala et al. (2014)       |
|                              | ANN              | –         | CO$_2$ content, meteorological parameters, etc. [12 input parameters]          | Superficial gas flux            | Error variance: 36          | Scozzari (2008)                 |
| Landfill area estimation     | ANN              | 180       | Trip number of vehicles, month of the year [2 input parameters]               | Waste quantity                  | $R$: 0.86                   | Hoque and Rahman (2020)         |
|                              | ANFIS            | –         | Age groups of 0–14, 15–64 and 65+ [3 input parameters]                         | Waste generation rate           | RMSE: 3.860                 | Younes et al. (2016)           |
| Landfill surface temperature | ANN              | 7830      | Meteorological parameters [6 input parameters]                                | Landfill surface temperature    | $R$: 0.884                  | Abu Qdais and Shatnawi (2019)   |
addition, extensive monitoring data on waste generation are increasingly collected, and some useful data are available for free usage (Niska and Serkkola, 2018). In the future, it will be necessary to integrate existing data resources, establish a monitoring data management platform, and a unified data standard specification for multiple formats and types of monitoring data from remote sensing, surveying and mapping, geographic information, aerial photography, Internet of Things, and radio frequency identification (RFID). Data fusion technology can be used to realize data interconnections among different information systems, databases, and data types. In addition, it is necessary to establish and improve the mechanism of data opening and sharing, thereby creating a good data opening environment.

Data pre-processing and feature selection

Data pre-processing is a very important step in ML, and whether the data are processed properly has a great impact on the training and prediction results. Missing values and noise were prevalent in the current study. Linear interpolation or mean value substitution is a common solution for data completion (Birgen et al., 2021; Cubillos, 2020; Dissanayaka and Vasanthapriyan, 2019). However, these methods can easily lose information. Recently, Fallah et al. (2020) reconstructed missing data in CH₄ generation rate records using ANN and reduced the MSE greatly, which provides a good way for the process of MSW data missing in the future.

Moreover, appropriate selection of the input variables can influence the performance of the learning process and the robustness of ML models. The proper selection would enhance the performance and decrease the complexity of the modeling. The MSWM process is complex and involves a large number of influencing factors and operating control parameters (Norhayati and Rashid, 2018; Soto-Paz et al., 2020). There is also a problem of multicollinearity when analyzing the importance of features (Kannangara et al., 2018; Shaha et al., 2020). The highly correlated features not only make it difficult to interpret the importance of variables, but also reduce the performance of the model. In the current studies, only a few of them have carried out feature selection, mainly using the Pearson correlation coefficient (Hoang et al., 2019b; Wu et al., 2020a), but the main disadvantage of this method is the simple use of a linear function to extract the relevant information between the input and the target. Hence, more advanced methods, such as mutual information and Lasso, are recommended for this purpose.

ML algorithm selection

There are currently a variety of ML algorithms, but there is no clear standard for which algorithm should be chosen. One solution is to try different algorithms on the same dataset to obtain the best model for the target case. According to existing literature, ANN is the most commonly used algorithm, whether as the main model or as a comparable one for evaluation, but ANN is prone to overfitting and is a black box. In recent years, the application of ensemble learning (e.g. gradient boosting) and DL models in MSWM has gradually increased (Akanbi et al., 2020; Johnson et al., 2017; Kontokosta et al., 2018). Gradient boosting improves prediction accuracy by combining the outputs of weak models to form a single consensus model (Adoogha et al., 2019). DL can design features automatically, which can improve system performance. For example, CNN has been widely used in MSW classification with high accuracy (Togacar et al., 2020). However, DL methods are rarely used in other MSWM processes. Therefore, it is necessary to develop new ML models. DL models based on an RNN or LSTM can be established to predict gas emissions. Meanwhile, to quickly respond to complex environmental changes, reinforcement learning can also be used for real-time online route optimization for waste collection and transportation (Zhou et al., 2020).

Model comparison

An important step in evaluating effectiveness of ML models is to conduct comparison experiments. Many of the current studies use only one ML algorithm (e.g. ANN) without comparison with other models (Duan et al., 2020; Kulisz and Kujawska, 2020). Some other studies only compared the results of different algorithms (e.g. ANN, SVM, etc.) to select the best one (Abbasi and El Hananadeh, 2016; Solano Meza et al., 2019), lacking comparisons between different papers. Although these studies claimed to have good results, the lack of comparisons with similar methods makes it difficult to be convincing. Only a few studies used the results in other papers as a comparison baseline directly, but due to the different datasets used, such a comparison is not convincing (Abushammala et al., 2014; Bagheri et al., 2019; Kannangara et al., 2018; Togacar et al., 2020). Recently, Mao et al. (2020) compared the performance of different CNN models proposed by several studies based on the same dataset of TrashNet, proving that their model can achieve better accuracy. It is also worth noting that most of the current research does not have a detailed model description, that is, the fine-tuning description of the model, such as the setting of hyperparameters, so their results are often difficult to reproduce. In the future, it is important to develop a baseline and produce transparent descriptions of ML models, which are conducive to the continuation of scientific research in the previous findings and the rapid development of applications in this field.

Model interpretability

For the application of ML in MSWM, we need to find the hidden knowledge behind the model, formulate policies, and assist decision-making based on the results. Despite the good performance of the ML algorithms mentioned above (e.g. ANN, SVM, and ANFIS), they are considered as black boxes (Hoque and Rahman, 2020). It is very difficult to gain a clear and insightful
interpolation of these models owing to the sophisticated relationship between the variables and the output. However, in the past few years, there has been a lot of work in computer science to make ML models interpretable (Carvalho et al., 2019), but in the field of MSWM, such research is still rare. Recently, Shaha & et al. (2020) used RF and DL methods to monitor and predict the operational performance of the boiler used in waste incineration plants, and introduced feature importance, partial dependence plots, and LIME to interpret the reason behind the model’s output, which can help the operator make reliable decisions. Interpretability can be used by scientists to understand the reason behind the model’s output and explain their models to stakeholders, so they can understand the value of the findings and make informed decisions. This will be an important direction for the application of ML algorithms in MSWM.

Conclusion

We reviewed 226 studies that used ML algorithms for MSWM. The data-driven approach can effectively model the complex relationship in MSWM, attracting the attention of researchers from both developed and developing countries (Abunama et al., 2019; Adeogba et al., 2019; Edgar et al., 2020; Shaha & et al., 2020), and becomes the foundation of refined management, providing a basis for managers for planning and decision making. The main conclusions are as follows:

First, ML algorithms have been applied in all aspects of MSWM, from waste generation to sorting and transportation, to final treatment and disposal, including waste generation prediction, bin level detection, route optimization, waste classification, forecast of compost maturity, HV and landfill leachate, pollutant emission monitoring, process control, and optimization.

Second, there is increasing interest in assessing MSWM with ML models, especially in the last two years, and the number of articles in this field has grown rapidly. ANNs are widely implemented models, including different algorithms such as feed forward, back propagation, radical basis function (RBF), autoregressive, and more than half of the literature reviewed used ANN as the main or comparison model. Recently, the application of DL technology in MSWM has gradually increased, mainly focusing on the use of CNN for waste classification, and there are also a small number of research papers using LSTM to forecast waste generation (Cubillos, 2020; Huang et al., 2020b).

Third, due to the complexity of MSWM, the mature application and practices of ML methods still require a long time. At present, there have been practical applications such as intelligent garbage identification and classification. For example, an Android app, called SpotGarbage, employed a CNN to automatically detect and localize garbage (Mittal et al., 2016). ZenRobotics Ltd. designed a waste sorting robot incorporating DL technology, which has been used in many countries. However, most of the reviewed studies are exploratory, and it is expected that there will be more practical applications in the future.

Finally, to completely realize the role of ML algorithms, it is necessary to integrate existing data resources, break data islands and data barriers, realize data sharing, and achieve information acquisition and supervision in the whole process of waste disposal, so as to improve the timeliness and accuracy of waste management, turning the waste treatment process into highly intelligent operational and management activities.
Wu H, Tao F and Yang B (2020b) Optimization of vehicle routing for waste collection and transportation. *International Journal of Environmental Research and Public Health* 17: 4963.

Xue W, Hu X, Wei Z, et al. (2019) A fast and easy method for predicting agricultural waste compost maturity by image-based deep learning. *Bioresource Technology* 290: 121761.

Yang M and Thung G (2016) TranshiNet. Available at: https://github.com/garythung/trashnet (accessed 19 January 2021).

Ye Z, Yang J, Zhong N, et al. (2020) Tackling environmental challenges in pollution controls using artificial intelligence: A review. *Science of the Total Environment* 699: 134279.

Yi S, Yoo K and Hanaki K (2011) Characteristics of MSW and heat energy recovery between residential and commercial areas in Seoul. *Waste Management* 31: 595–602.

You H, Ma Z, Tang Y, et al. (2017) Comparison of ANN (MLP), ANFIS, SVM, and RF models for the online classification of heating value of burning municipal solid waste in circulating fluidized bed incinerators. *Waste Management* 68: 186–197.

Younes MK, Nopiah ZM, Basri NEA, et al. (2016) Landfill area estimation based on integrated waste disposal options and solid waste forecasting using modified ANFIS model. *Waste Management* 55: 3–11.

Zhang J, Petersen SD, Radivojevic T, et al. (2020) Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. *Nature Communications* 11: 1–13.

Zhou B, Song Q, Zhao Z, et al. (2020) A reinforcement learning scheme for the equilibrium of the in-vehicle route choice problem based on congestion game. *Applied Mathematics and Computation* 371: 124895.