REVIEW ARTICLE   **Open Access**

Intelligent control of nanoparticle synthesis on microfluidic chips with machine learning

Xueye Chen [iD][1] and Honglin Lv[1]

Abstract

Nanoparticles play irreplaceable roles in optoelectronic sensing, medical therapy, material science, and chemistry due to their unique properties. There are many synthetic pathways used for the preparation of nanoparticles, and different synthetic pathways can produce nanoparticles with different properties. Therefore, it is crucial to control the properties of nanoparticles precisely to impart the desired functions. In general, the properties of nanoparticles are influenced by their sizes and morphologies. Current technology for the preparation of nanoparticles on microfluidic chips requires repeated experimental debugging and significant resources to synthesize nanoparticles with precisely the desired properties. Machine learning-assisted synthesis of nanoparticles is a sensible choice for addressing this challenge. In this paper, we review many recent studies on syntheses of nanoparticles assisted by machine learning. Moreover, we describe the working steps of machine learning, the main algorithms, and the main ways to obtain datasets. Finally, we discuss the current problems of this research and provide an outlook.

Introduction

Nanotechnology is a science in which nanoparticles with sizes ranging from 1 to 100 nm are prepared through multiple synthetic pathways and nanoparticle structure and size are studied[1–4]. Nanoparticles have unique properties based on their nanoscale sizes that larger particles do not possess. Today, nanoparticles play irreplaceable roles in many fields, such as molecular biology, electronic sensing, organic and inorganic chemistry, medical therapy, and materials science[5–8]. With the rapid development of microfluidics, the preparation of nanoparticles in microreactors has received wide attention[9–11]. Microreactors have the advantages of better heat transfer capability, less reagent consumption, easy operation, and easy integration[12–15]. In the process of preparing nanoparticles, temperature conditions usually must be triggered to drive the reaction, and the better heat transfer capabilities of microreactors can make the reactions happen quickly and controllably[16,17]. These advantages make the use of microfluidics for nanoparticle preparation

a superior prospect for development. Since the properties of nanoparticles play different roles depending on their sizes and morphologies, syntheses of nanoparticles with the desired properties require precise control of size and morphology[18–20]. In a recent study, Zhang et al.[21] found that the average particle size and standard deviation for synthesized Cu_2O nanoparticles were 0.967 and 0.12 μm, respectively, when the voltage was 5 V. When the voltage was 8 V, the average particle size and the standard deviation decreased to 0.680 and 0.09 μm, respectively. Therefore, the sizes of nanoparticles can be adjusted with reaction conditions. Nanoparticles of different sizes exhibiting different properties in electron optics were obtained. Lundqvist et al.[22] studied six different examples of polystyrene nanoparticles made with three different surface chemicals (plain PS, carboxyl-modified, and amine-modified) and two sizes (50 and 100 nm). They found that different nanoparticle surface properties and sizes affected detailed protein corpuscles differently. The extent of changes observed in the natures of bioactive proteins in the corona varies when different surface chemicals are used. Different sizes of nanoparticles have different effects on protein coronas. However, with

Correspondence: Xueye Chen (xueye_chen@126.com)
[1]College of Transportation, Ludong University, Yantai, Shandong 264025, China

current technology, syntheses of nanoparticles with expected properties require large numbers of experiments and iterative tuning. This is costly in terms of time and increases the amount of reagents consumed. Therefore, precise synthesis of nanoparticles is currently the most daunting problem for researchers[23,24].

The most important problem that can be solved in the field of machine learning is prediction. Machine learning can construct complex relationships relating to data samples[25]. Machine learning is a branch of artificial intelligence that uses large numbers of known datasets to predict or respond to unknown data[26]. The model is constructed through algorithms that include both simple linear regressions and very complex nonlinear regressions, which ultimately allow the model to produce the best predictions[27–29]. Thus, the use of machine learning-assisted syntheses of nanoparticles can overcome the challenges of preparing nanoparticles with functionality. In a recent study, Ban et al.[30] performed predictive analyses with machine learning models, and protein predictions showed an R^2 value exceeding 0.75. This study provided a novel approach for accurately and quantitatively predicting the functional compositions of protein corpuscles that determine cell recognition and nanotoxicity with machine learning and thereby overcame the uncertainty involved in synthetic processes and provided precise syntheses of various nanoparticles.

In this paper, we first review the advantages of synthesizing nanoparticles on microfluidic chips with different types of microreactors. Second, we describe the working steps of machine learning, the main algorithms, and the main ways to obtain datasets. Next, we review many recent reports of machine learning-assisted syntheses of nanoparticles with desired properties. Finally, we discuss the current problems of machine learning and the future prospects.

Synthesis of nanoparticles by microfluidic technology

A microfluidic chip is a device that uses micron-scale channels to handle fluids[31–33]. It can handle micron- and nanoscale fluids mainly by taking advantage of its small size and thus greatly reduces sample consumption, reduces bioassay costs, and increases reaction rates. It also has the advantages of high integration and ease of operation[34–36]. It shows great promise for biochemical syntheses, drug development, medical diagnostics, microbial research, and nanoparticle preparation. The microreactor is also known as a microchannel reactor. It can be subdivided into micromixers, microexchangers, and microreactors[37]. It is a microreactor with feature sizes between 10 and 300 microns (or 1000 microns), and it is manufactured by using precision machining techniques. The "micro" in the microreactor indicates that the

channels of the process fluid are at the micron level, rather than referring to the small form factor of the microreactor or the small yield of product. Microreactors can mix reagents quickly and uniformly, and they can accelerate reactions by placing obstacles inside the microchannels to facilitate fluid mixing or by applying an external energy field. Since microreactors exhibit small sizes, they have the advantage of fast heat transfer; this improves the processes used for preparing nanoparticles, which require certain temperatures for the reactions to occur, by improving the heating rate and the reaction rate[38–40]. In current biochemistry, electronic sensing, and other fields, nanoparticles have irreplaceable advantages due to their unique properties. Therefore, research on the preparation of nanoparticles is crucial[41].

In traditional reaction engineering, the main methods used for preparing nanoparticles include evaporation coalescence, mechanical comminution, chemical reduction, chemical deposition, and sol-gel methods[42–45]. The first two methods are commonly used in the early stage. Evaporation coalescence is a physical method that involves heating metal materials under a high vacuum, evaporating them into molecules or atoms, and then coalescing them into nanoparticles[46,47]. The advantage of this method is that the nanoparticles produced are of high purity and the particle sizes are easy to control, but the environmental conditions are demanding, and the operations are difficult. Mechanical comminution is also a physical method that works by splitting a solid mass through impact from a comminution force to produce finer particles[48,49]. Common crushing forces include shearing, crushing, impact crushing, and grinding. The chemical reduction method is a common method for preparing nanoparticles. According to the state of the reducing agent used, the reactions can be divided between liquid-phase and gas-phase reduction methods[50–54]. The liquid-phase reduction method is a process in which a metal salt solution is directly reduced by the reducing agent at room temperature and atmospheric pressure to prepare nanoparticles. However, in the traditional method of preparing nanoparticles, it is difficult to control the sizes of nanoparticles precisely, and the consumption of reaction raw materials during the preparation process is also large. However, compared with traditional reactor processes, microreactor technology has many advantages, such as high mass and heat transfer efficiencies, short reaction times, no amplification effect, safety and reliability, high integration, and green production processes[55–57]. Microreactor technology can significantly enhance the reaction process; it constitutes an innovation in chemical synthesis, and it provides an efficient and convenient operating platform for chemical production. Therefore, the preparation of nanoparticles on microfluidic chips is imperative[58,59].
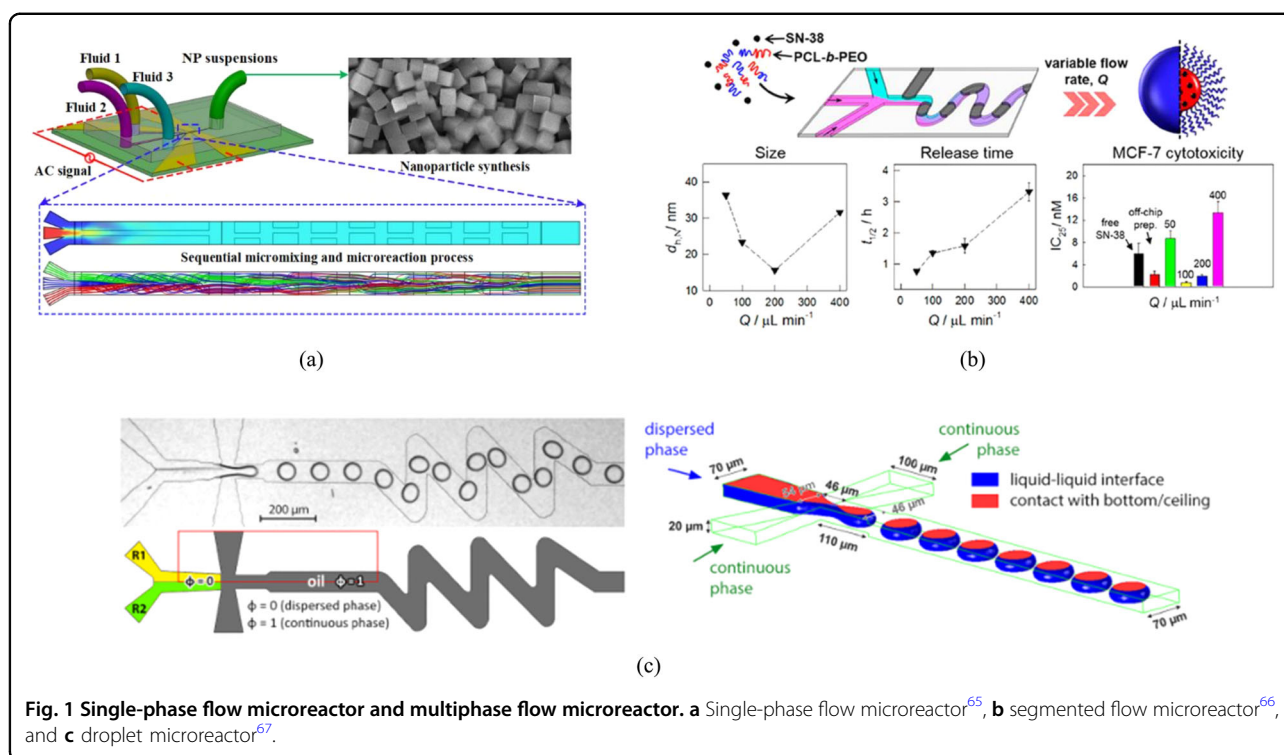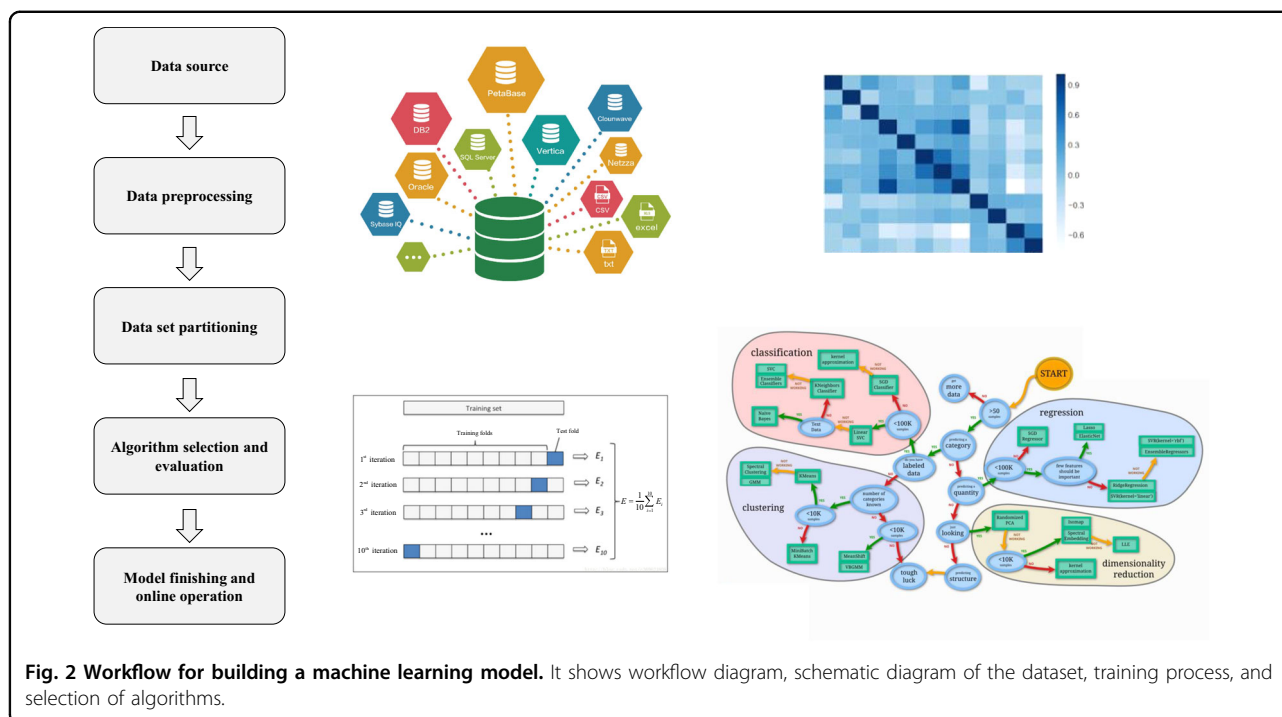
Fig. 1 Single-phase flow microreactor and multiphase flow microreactor. a Single-phase flow microreactor[65], **b** segmented flow microreactor[66], and **c** droplet microreactor[67].

There are primarily two different types of microreactors used for preparing nanoparticles on microfluidic chips, namely, single-phase continuous flow microreactors and multiphase flow microreactors[60–64]. Single-phase continuous flow microreactors (Fig. 1a) use continuous laminar flow inverse mixing and reaction of multiple fluids within the microchannel. To ensure adequate mixing of the reactants, geometric baffles can be placed within the microchannels or the geometries of the microchannels can be changed. This facilitates higher reaction rates and allows full conversion of reactants to products. Sun et al.[65] developed a unique platform for sequential micromixing-assisted nanoparticle syntheses by using AC electric heat flow. Sequential micromixing was achieved by flexibly adjusting the volume of each fluid and the AC voltage inside the three-fluid mixer. Nanoparticles of cobalt-iron Prussian blue analogs were synthesized by using two-fluid mixing and three-fluid mixing methods.

Multiphase flow microreactors can be divided into segmented flow microreactors (Fig. 1b) and droplet microreactors (Fig. 1c)[66,67]. They are characterized by multiple inlets, and a gas and a liquid or a liquid and a liquid are introduced separately, thus forming a continuous and a dispersive section. In the dispersion section, the various reaction reagents are fully mixed, and chemical reactions take place. The advantage of the multiphase flow reactor is that the reaction zone is not in contact with the outer channel walls. This greatly reduces the risk of contaminating the prepared products and reduces the risk of clogging the microchannels. However, the system is not easy to manipulate, and there is a chance of reagent dispersion. It is also not easy to add new reagents for continuous reaction experiments. Kumar et al.[68] segmented flow in a spiral poly(methyl methacrylate) microreactor by using silver nanoparticles reduced/sealed with stearic acid sophorolipids as a model system with kerosene and air as inert phases. The segment plug size and slip velocity controlled the nature of mixing within the segment plug of the reactant phase, and they controlled the size distribution of the nanoparticles.

Current research on nanoparticles reveals that different sizes of nanoparticles have different electronic, sensing, and optical properties. The factors that usually affect the sizes of nanoparticles are reaction conditions and the type of reactants used, including the concentration ratio of the reactants, reaction temperature, reaction time, type of reaction reagents, reaction precursors, and the structures of the microchannels. All of these factors can have different effects on the particle sizes and surface properties of the synthesized nanoparticles. The synthesis of functional nanoparticles is a problem that researchers have been working to overcome. Machine learning can be used to construct relationships among predictive variables and target properties, use reaction conditions to determine the properties of nanoparticles, precisely control the syntheses of nanoparticles with desired functionalities and thus

Fig. 2 Workflow for building a machine learning model. It shows workflow diagram, schematic diagram of the dataset, training process, and selection of algorithms.

enable accurate predictions of nanoparticle properties and applications[69,70]. In this paper, we review the extensive literature on machine learning-assisted preparation of nanoparticles to facilitate the reader's understanding of research reported in recent years.

Machine learning

In this section, we describe the workflow of machine learning (as shown in Fig. 2), the main algorithms contained, and the main processes used to access datasets.

The machine learning process
Data sources

The first and most important step in machine learning is the preparation of a dataset, which is crucial. The size and quality of the dataset directly affects the accuracy of the machine learning model[71]. A dataset comprises a large set of independent variables and target properties. The way a dataset is obtained will be reviewed in detail in a later section. The dataset requires the identification of independent and dependent variables. In nanoparticle preparation techniques, the independent variables are generally selected as reaction conditions (e.g., reaction temperature, reaction time, concentration ratio of chemical reagents, reaction precursors, reaction ligands, type of solution reagents, channel structure of the microreactor, and external stimuli), while the dependent variables are generally the sizes and shapes of the nanoparticles as well as the electronic and optical properties[72].

Data preprocessing

Once the dataset is constructed, it needs to be pre-processed to ensure high quality. Since large amounts of data obtained directly from numerical simulations or experiments may have errors and large gaps, data pre-processing is performed[73]. Data preprocessing is actually data cleaning, data sorting or general data processing. It refers to the process of checking and improving data in various ways to correct missing values or spelling errors, normalizing or standardizing values to make them comparable, converting data, and other issues. Larger datasets provide more features that can be used for training machine learning models, which will lead to much higher predictive power[74].

Dataset division

There are two approaches used to partition datasets. The first approach is to divide the dataset into a training set and a test set, which generally exhibit a ratio of 7:3[75]. In the machine learning model development process, it is expected that the trained model will perform well with new, unseen data. To make machine learning models perform well on more than just the training set, it is necessary to divide the dataset into two parts, one for training the machine learning model and the other for testing the predictive power. Another approach is to divide the dataset into a training set, validation set, and test set, which generally have a ratio of 6:2:2[76]. The training set is used to build the machine learning model, and the validation set is used to evaluate the model.

Predictions are made accordingly, and the model with the best predictive power is selected based on the results obtained with the validation set. The validation set operates in a similar way to the training set. Importantly, the test set is not involved in building and preparing the machine learning model; it contains a separate set of samples that are set aside during the training of the machine learning model to adjust the hyperparameters of the model and to make a preliminary assessment of the model's capabilities.

Selection and evaluation of machine learning algorithms

Machine learning has a rich set of algorithms to choose from for different tasks. Machine learning algorithms can be classified into supervised learning, unsupervised learning, and reinforcement learning. In a later section, we will elaborate on the three types of algorithms[77]. For making the optimal choice of algorithms, it is best to test the mainstream algorithms separately, evaluate the results they provide, and select the one with the best performance. However, each algorithm has different characteristics and is suited for different tasks. However, for the general selection of algorithms, we can also use general linear criteria for selecting based on the different characteristics of the algorithms. For example, if the training set is small, a high bias/low variance classifier (e.g., a plain Bayesian classifier) is preferred to a low bias/high variance classifier (e.g., a k-nearest neighbor classifier) because the latter is prone to overfitting. However, as the size of the training set is increased, low bias/high variance classifiers start to perform more effectively (they have lower asymptotic errors) because high bias classifiers are not sufficient to provide accurate models[78].

Model finishing and online operation

After training of the machine learning model is completed through the above process, the trained model must be organized into files to ensure easy use and smooth operation[79]. The files used are the model file, coding file, metadata file (algorithm, parameters, and results), and dataset file (a CSV dataset file of independent and dependent variables). Finally, the trained machine learning model is put online and run[80].

Datasets and algorithms
Main access to the dataset

As described above, a dataset must be prepared before building a machine learning model. The dataset generally consists of a training set and a test set[81,82]. The dataset used to train the machine learning model is called the training set, and the dataset used to verify the accuracy of the machine learning model is called the test set. Generally, they have a ratio of 7:3. The predictive power is usually influenced by the amount of data in the test set,

the choice of kernel function, and the definition of the loss function. The larger the dataset is, the more accurate the predictions of the machine learning model, but large datasets may substantially increase the computational cost and lead to long computation times[83,84]. Too much data may also cause overfitting. Before making predictions, machine learning models need to be validated to evaluate the two main ways of constructing datasets. One is to generate the results through numerical simulations or new experiments, which are processed to form a dataset. This is the best way to build a dataset that clearly represents the relationships among the desired variables and the target. However, the time cost for numerical simulations or experiments is too high when the datasets are too large[85,86]. The second method is to sample the datasets described in published articles by reviewing the relevant literature. However, since the variables and targets of the constructed dataset may differ depending on the purpose of the work, a dataset thus acquired cannot be used directly and must be processed, for example, by removing some variables or redundant data. This method for dataset acquisition is efficient if one wants to optimize the parameters of the current algorithmic model to achieve the same purpose as the prior work[87]. Datasets obtained online can be utilized directly to build machine learning models and optimize the parameters in those models to minimize losses. This is also a research method. In addition, there exist a number of open source dataset sites on the web that contain rich databases[88]. This is a convenient way for novice machine learners to learn how to train their models. In this paper, we have summarized some open source databases for readers to use (shown in Table 1).

When the acquired dataset has missing values or outliers, we can process it by ensuring that the average value of data close to the missing data replaces the missing value or the missing values are replaced according to a statistical model. Before building the machine learning model, we can normalize the data in the prepared dataset, which can improve the accuracy and computational convergence velocity of the machine learning model.

Algorithms for machine learning

Machine learning algorithms are usually classified into three types: supervised learning, unsupervised learning, semisupervised learning, and reinforcement learning[89,90].

Supervised/unsupervised learning In machine learning, the main difference between supervised and unsupervised learning is the presence or absence of human supervision. In supervised learning, the training samples are used with labels. The training set is composed of input variables and output targets[91,92]. Supervised learning can generally be used to predict data and to classify data samples. The
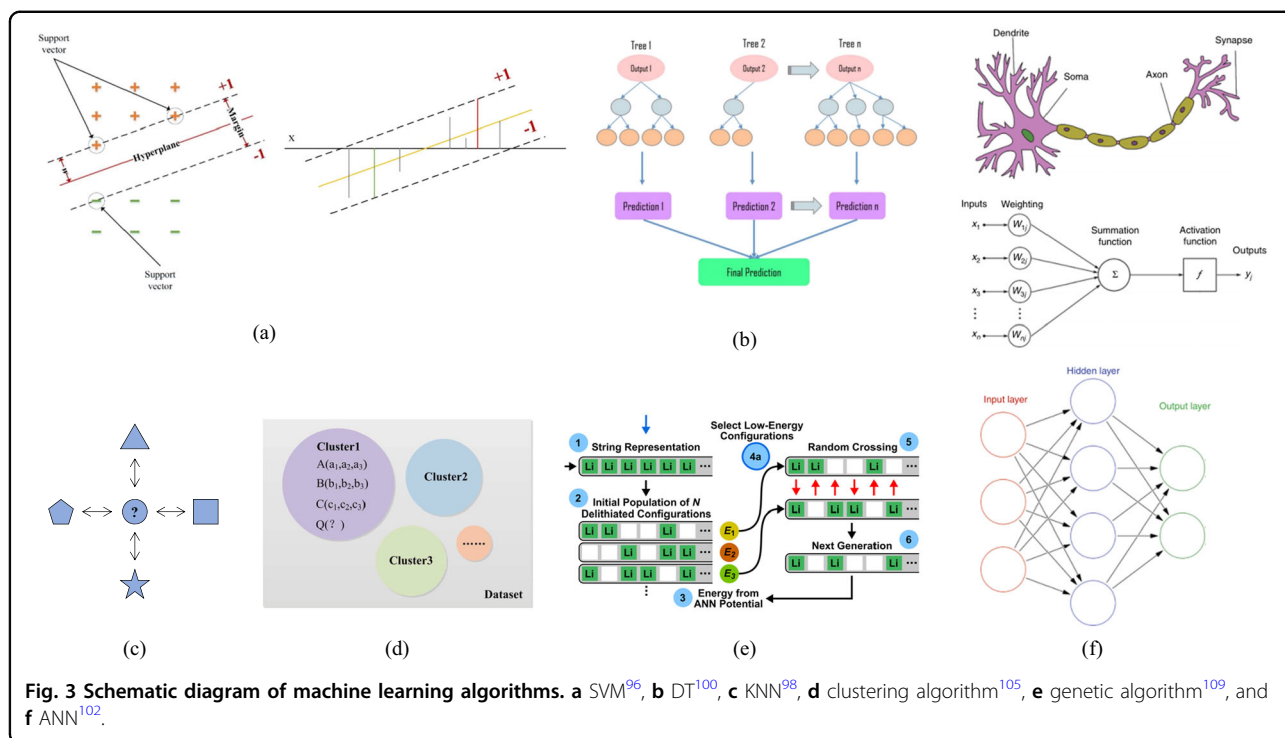
Table 1 Access to the dataset.

Database	Description	Web address
The Inorganic Crystal Structure Database[124]	Experimental characterization data of inorganic crystal structure	https://icsd.fiz-karlsruhe.de/index.xhtml
Materials Project[124]	Calculation of data of properties of known and hypothetical materials	https://materialsproject.org
Awesome public datasets	It is a list of a series of datasets categorized by topic and publicly maintained by the community, such as biology, economics, education, etc.	https://github.com/awesomedata/awesome-public-datasets
Materials Platform for Data Science[124]	Peer-reviewed crystal structure, phase diagram, or physical property	https://mpds.io/#modal/menu
Materiae[124]	Topological material database	http://materiae.iphy.ac.cn/
Kaggle dataset	The website contains a large number of real datasets of various shapes, sizes, and formats	http://www.kaggle.com/datasets
Droplet[87]	Datasets on droplet generation with DI water and mineral oil	http://dafdcad.org/download
Droplet1[87]	Datasets on droplet generation with LB bacterial cell media and mineral oil	http://dafdcad.org/download
Droplet2[87]	Datasets on droplet generation with DI water and light mineral oil	http://dafdcad.org/download
Google's dataset search engine	It is a toolbox that can search datasets by name	https://toolbox.google.com/datasetsearch

difference between regression and classification algorithms is the type of output variable; quantitative output is called regression, or continuous variable prediction, and qualitative output is called classification, or discrete variable prediction[93]. Of these, predictive models are more widely used, and in this work, accurate prediction of the functional properties of nanoparticles is the first priority. There is a nonlinear link between reaction conditions and the properties of nanoparticles. If we can predict the reaction conditions of nanoparticles based on the desired properties of nanoparticles by using predictive models, then we can achieve precise control of nanoparticle syntheses and functional nanoparticles. That is why regression models for machine learning are highlighted and reviewed. Regression modeling algorithms for supervised learning are divided into the following categories[94,95] (Fig. 3).

Support vector machine algorithm (SVM): in the regression model algorithm, it is assumed that the sample set of the dataset is $D = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$, and the predicted regression model is $f(x) = \omega^T x + b$, where b are the parameters of the regression model. The essence of the regression problem is to find a straight line or curve so that it fits the data points perfectly[96]. For a sample dataset in a traditional regression model, the loss is defined by the difference between the predicted value $f(x)$ and the true value y. The loss is minimized when these two values are equal. However, for support vector machine regression models (SVR), the principle is slightly

different from the traditional approach. It establishes a nonlinear relationship between the variables and the target by first defining a constant ε. For some sample data points in the dataset, if $|f(x) - y| < \varepsilon$, there is no loss at all. If $|f(x) - y| > \varepsilon$, then a loss is defined to exist, and the loss is $|f(x) - y| - \varepsilon$. For different practical problems of constructing nonlinear regression models, many kernel functions in SVR algorithms are offered, including linear kernel functions, polynomial kernel functions, Gaussian kernel functions, Laplace kernel functions, and sigmoid kernel functions. However, kernel functions have no general selection criterion, so it is difficult to choose an exact kernel function[97].

K-nearest neighbor algorithm (KNN): the KNN algorithm can also be used to address regression problems. It is based on a prediction principle of finding the k-nearest neighbors of a target sample and assigning an average value for a target attribute of these neighbors to that sample or by using the weighted average method, in which the weight is set by the distance between the neighbor and the target sample, and the distance is inversely proportional to the weight[98]. The value of the corresponding attribute of that sample can be obtained by the above calculation, which leads to a prediction for which k is the algorithm parameter. It is determined by the number of samples in the dataset. The larger the value of k is, the larger the deviation of the model and the less sensitive it is to noisy data. A large value of k may cause underfitting. The smaller the value of k is, the larger the variance of the

Fig. 3 Schematic diagram of machine learning algorithms. a SVM[96], **b** DT[100], **c** KNN[98], **d** clustering algorithm[105], **e** genetic algorithm[109], and **f** ANN[102].

model. When the value of k is too small, overfitting may result. Therefore, the choice of hyperparameter k also needs to be debugged. However, KNN has the advantages of simplicity and speed when training samples[99].

Decision tree (DT): the DT model for solving regression problems is known as a regression tree[100]. The principle of the regression tree is to divide the variables in the dataset by selecting multiple cut points. Loss functions are calculated for both parts of the divided data, and the cut point is compared with the cut point that results in the smallest loss result. The loss function is defined as follows:

$$\text{Loss}(y, f(x)) = (f(x) - y)^2 \tag{1}$$

The dataset is divided into two parts by the optimal cut point. The above operation is continued for these two parts of the dataset, the optimal cut is selected step by step, and the regression tree model is finally constructed. The result of the regression tree is actually a segmentation function. When using a regression tree for prediction, training and prediction are fast, and it is easy to obtain nonlinear relationships between variables and targets, but there are problems such as inaccurate prediction accuracy and easy overfitting. Regression trees are rarely used to construct prediction models in real engineering problems[101].

Artificial neural network (ANN): ANN refers to a complex network structure formed by a large number of processing units, e.g., neurons, connected to each other. It works well in constructing nonlinear relationships

between variables and targets in a dataset[102]. Each neuron has a weight, which means that the input value sent to each neuron is multiplied by this factor. There is an activation function in the neuron so that the output calculation will be done by the combination of multiple nonlinear functions. The ANN model is constructed with an input layer, a hidden layer, and an output layer. One of the neuron layers constitutes a combination of multiple neurons. The number of neurons can be set flexibly, and losses of the prediction model can be minimized by adjusting the number of neurons and the parameters in the activation function to achieve accurate predictions[103]. ANNs have powerful fitting capabilities. They can approximate arbitrarily complex nonlinear relationships very well and perform self-learning and self-adaptive tasks for uncertain systems. In addition, they are robust and exhibit fault tolerance. In general, their predictive abilities and computational costs are superior to those of other algorithms[104].

The various algorithms listed above for regression prediction demonstrate that machine learning can assist in the preparation of nanoparticles. The powerful ability of machine learning to fit nonlinear relationships has led to widespread interest in its use for easy preparation of nanoparticles with functionalities.

The samples in unsupervised learning data are not labeled. The classes of the samples are not known during the training process, and training samples are needed to determine the similarities of samples in the dataset so

similar samples can be clustered into one class. In a nutshell, this is clustering. The methods used in clustering are listed below[105–107].

(1) K-means clustering, which is based on dividing all data samples into K mutually exclusive groups. Determination of the size of K is to be considered.

(2) Hierarchical clustering, which is based on the principle of dividing all data samples into groups and subgroups, thus forming a tree similar to those used in genealogy.

(3) Probabilistic clustering classifies all data samples according to their probabilities, where the probability can only be 0 or 1. K-means clustering is a special form of this. Probabilistic clustering is also known as "fuzzy K-means".

In experimental syntheses of functional nanoparticles on microfluidic chips, researchers mainly pursue the construction of models that can predict the properties of nanoparticles. Therefore, the clustering algorithm of unsupervised learning is not applicable. Unsupervised learning is used because of its powerful clustering ability, and it can cluster unconnected massive data samples into several classes based on their inherent similarities. It is commonly used to process images, and based on its properties, we conclude with the thought that if it can be applied in the field of nanoparticle synthesis, it will certainly bring breakthroughs in areas such as chemical synthesis[108].

Reinforcement of learning Reinforcement learning differs from supervised learning in that it emphasizes action based on the environment to achieve the maximum expected benefit, e.g., to achieve optimization. Unlike supervised learning, which has its own explicit goals, reinforcement learning can be seen as a system that keeps scores, remembers and uses actions to obtain low and high scores, and then continuously asks the machine to achieve high scores while avoiding low scores. It is applicable to a wide variety of tasks and captures many of the essential features of artificial intelligence, such as a sense of causality and uncertainty[109,110]. A key aspect of reinforcement learning is that it can be used as a substitute for learning good behavior. This means that it will gradually change or acquire new behaviors and skills.

In the process of achieving optimization, genetic algorithms or Bayesian optimization (BO) algorithms combined with algorithms in supervised learning can be used jointly to construct prediction models. For example, in the process of constructing a prediction model using ANN, the predictive accuracy of the obtained model differs based on the number of neurons and the settings of the model parameters. If the parameters of the prediction model can be optimized by combining optimization algorithms to minimize losses, the accuracy of the final prediction model will be improved by one level. In general, the fitness factor is optimized by defining the fitness function, or the function is optimized by defining the loss function. Their common purpose is to maximize the value of the predictive model by reinforcement learning, e.g., loss minimization.

Machine learning-assisted preparation of nanoparticles

Preparation of nanoparticles on microreactors has a large number of synthetic paths available, so it is difficult to achieve precise control of particle sizes and thus achieve syntheses of nanoparticles with the desired properties. The use of machine learning would be beneficial in terms of accurate syntheses of nanoparticles[111–113]. The unique optical properties of metallic and semiconductor nanoparticles allow them to be used in areas such as electronic sensing, solar cells, optoelectronic devices, and quantum information technology[114–116]. Here, we review machine learning-assisted syntheses of metal nanoparticles and inorganic semiconductor nanoparticles reported in recent years (Table 2).

Metal nanoparticles

Metal nanoparticles exhibit unique properties in various fields such as pharmaceuticals, energy, and catalysis. In particular, metal nanoparticles play vital roles in the fields of medicine and pharmaceutical science[117,118]. Among metal nanoparticles, gold and silver nanoparticles are the most important and frequently used. Moreover, nanoparticles with different sizes and morphologies have different properties and thus perform different functions. Therefore, when preparing nanoparticles for a specific application, precise control of nanoparticle size and morphology is critical. The use of machine learning-assisted synthesis can solve this problem[119,120].

In reactions using citrate and gold salts for preparation of nanoparticles, the concentration ratio of citrate/gold(-III), scanning velocity, and radiation intensity affect the sizes of gold nanoparticles. In this study, the nanoparticles were characterized by using optical, spectroscopic, and transmission electron microscopy. ANNs were used as the main models for prediction; they only need a large dataset and can establish complex relationships among inputs and outputs through powerful internal algorithmic capabilities. Since the sizes of nanoparticles can be indirectly characterized by calculating the local surface plasma absorption maxima of nanoparticles, ANN models were developed to predict the nanoparticle plasma absorption maxima, which allowed predictions of nanoparticle sizes. In constructing the neural network, the dataset consisted of three inputs, e.g., the citrate/gold(III) salt ratio, intensity, and scan rate, and one output, e.g., the absorption

Table 2 Machine learning-assisted synthesis of nanoparticles.

Nanoparticles	Description	Machine learning models	Ref.
Gold nanoparticles	The relationship between the ratio of sodium citrate to gold salt, intensity, scanning velocity, and absorption maxima was constructed by ANN to enable prediction.	ANN	[121]
Gold nanoparticles	Machine learning models (LASSO regression, RR, ENR, SVM, RF) were used to construct the relationship between the size of nanoparticles, the amount of reagents, and the type of bacteria on the antibacterial effect of nanoparticles.	LASSO regression, RR, ENR, SVM, RF	[122]
Silver nanoparticles	BO based on Gaussian processes is combined with DNN to optimize machine learning models and improve the accuracy of model predictions.	BO, DNN	[124]
Gold nanoparticles	The unique optoelectronic properties of gold nanoparticles are used as a driving factor to explore the conditions for synthesizing new shapes of nanoparticles based on genetic algorithms.	Genetic algorithm	[125]
Colloidal quantum dots	By constructing a neural network model with eight inputs, two hidden layers, and two outputs, the effects of Pb:S concentration and temperature ratio on half-peak half-width and band gap were predicted.	ANN	[131]
CdSe quantum dots	Predicting the properties of nanoparticles by building an integrated neural network.	ENN	[102]
CdSeTe quantum dots	The complex relationships between FWHM, maximum, intensity, and reaction conditions of quantum dots are constructed by a metamodeling algorithm using kriging, and then optimized.	Metamodeling by Krieger method	[132]
CdSe quantum dot	By defining the "dissatisfaction coefficient" and using the noise-resistant global search algorithm to iteratively update the response conditions, the desired optimization goal is finally achieved.	Noise-resistant global search algorithm	[133]
CsPbBr3 type perovskite	Linear regression, quadratic regression, quadratic regression with cross terms, and SVM regression models were constructed for prediction, respectively, and their predictive abilities were compared.	SVM	[127]
Inorganic perovskite	The quantum yield of inorganic chalcogenide nanoparticles and the composition at the target band gap are optimized by autonomous optimization with polydispersity.	Self-optimizing system	[138]
CsPbBr3 type perovskite	By using active learning algorithms, the machine learning model is made to explore the optimal synthesis route autonomously.	Active learning algorithm	[140]

maximum. A total of 118 data samples were included in the dataset. To improve the accuracy of the training model, the data in the dataset were preprocessed, and the input and output values were mapped to a range $[-1,1]$. The sigmoid function was chosen as the activation function for the neurons in the hidden layer. The neural network was comprised of a multilayer feedforward network and backpropagation learning algorithm. The weights and bias values were adjusted to minimize the mean square error (MSE) objective function. The MSE is defined by the following equation[121]:

$$MSE = \frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n} \qquad (2)$$

where x_i is the experimental value, y_i is the predicted value of the ANN model, and n is the number of samples in the dataset. Finally, a predictive model for nanoparticles was successfully constructed. The results showed that the absorption maximum and the associated gold nanoparticles sizes increased with decreasing scan intensity, decreasing citrate to gold(III) ratio, and increasing scan velocity (Fig. 4c–e). It was found that radiation intensity had the most important impact on the sizes of the gold nanoparticles, followed by the scanning velocity and the ratio of citrate to gold(III).

In the field of medical drugs, nanoparticles are currently used as alternatives to antibiotics because of their unique antibacterial abilities. The sizes of the nanoparticles, the amount of exposure agent used and the type of bacteria have a direct impact on the antibacterial effects shown by the nanoparticles. Machine learning models can be used to predict the antibacterial activities of nanoparticles. A

Fig. 4 Machine learning-assisted synthesis of metal nanoparticles. a Performance metrics[122], **b** random forest attribute importance analysis results[122], **c** prediction of gold nanoparticle size by ANN (scanning velocity)[121], **d** intensity[121], and **e** citrate:gold(III) ratio changes (c1 and c2)[121].
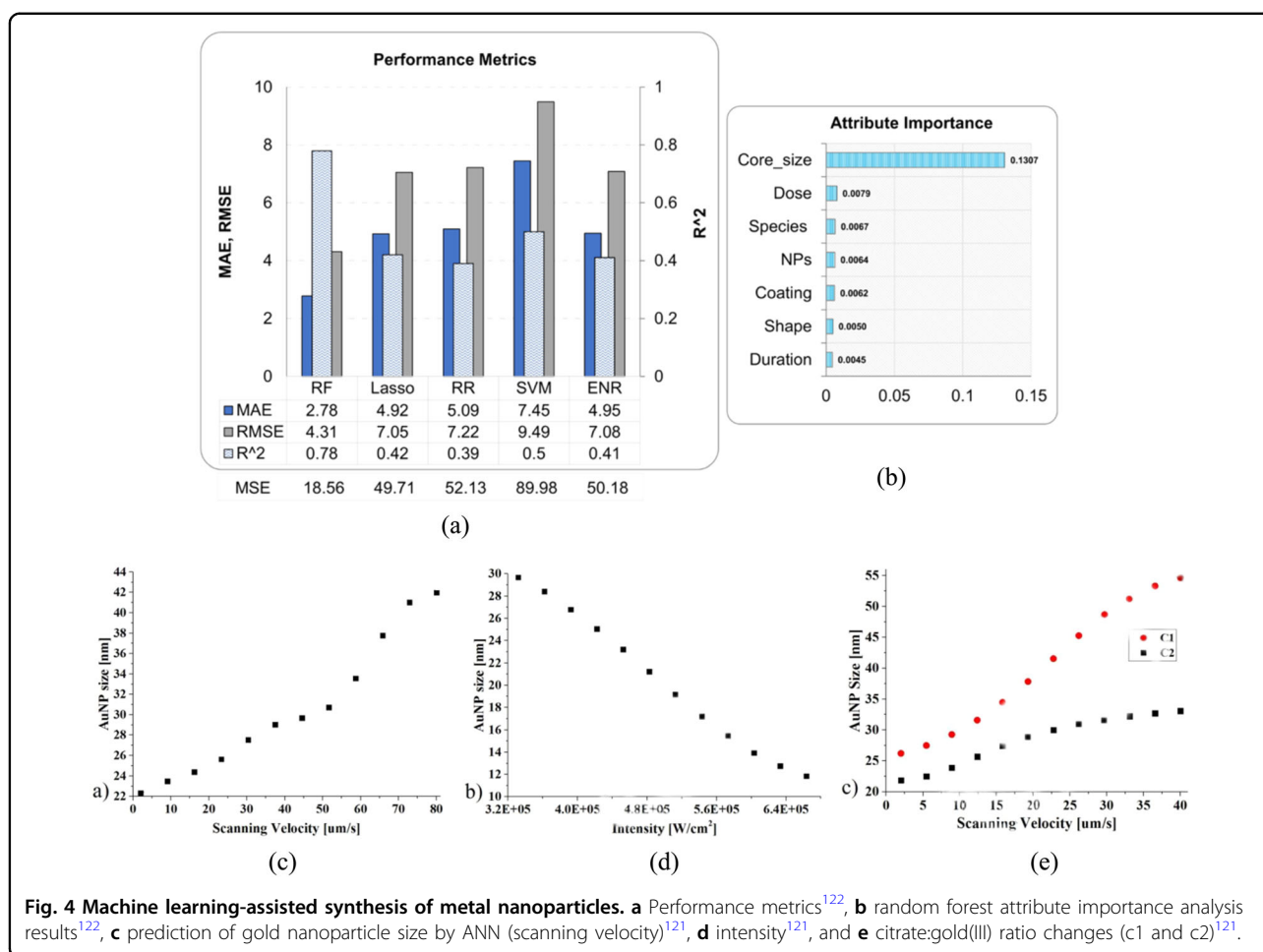
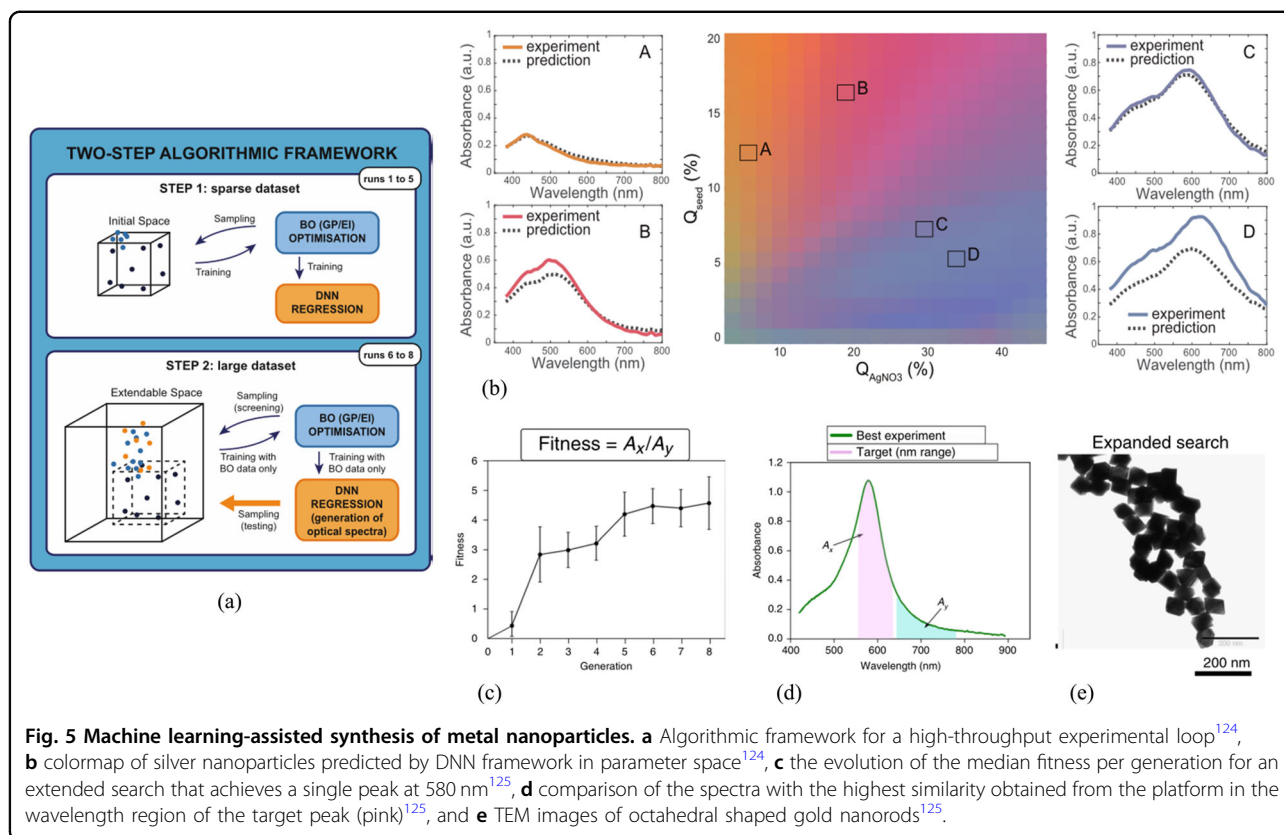dataset was obtained by compiling data from more than 60 papers in the literature. The data samples were homogenized and preprocessed before constructing the predictive model, which enhanced the quality of the data. In constructing the machine learning models, various regression algorithms were used to construct the predictive models to compare the predictive powers of different algorithms, and the constructed models were validated with different performance metrics. The model results obtained with the various algorithms in the figure clearly showed that the random forest (RF) model had the smallest error and the highest R^2 score with the least absolute shrinkage, selection operator (LASSO) regression, ridge regression (RR), elastic net regression (ENR), and SVM algorithms (Fig. 4a, b)[122]. The extents to which reaction conditions affected the antimicrobial properties of the nanoparticles were compared. Nanoparticle size was the most important property in determining the antibacterial effect. The amounts of reagent and bacterial species were relatively important, followed by the coating, shape, and duration. Syntheses of nanoparticles assisted

by machine learning help researchers synthesize useful and functional nanoparticles.

For nanoparticle syntheses, particle size is an important parameter. To investigate the factors affecting particle sizes, a microfluidic chip consisting of a Y-shaped mixing microchannel and two piezoelectric valveless micropumps was designed for the study[123]. The mixing and reaction rates of gold salt and reducing agent were adjusted by controlling the switching frequency of the piezo-point valveless micropump, which resulted in syntheses of nanoparticles of different sizes. The results showed that the sizes of the gold nanoparticles synthesized in this way increased as the mixing time decreased. This indicated that a machine learning model could be constructed to use the relationship between particle size and reaction time to control the sizes of the synthesized nanoparticles.
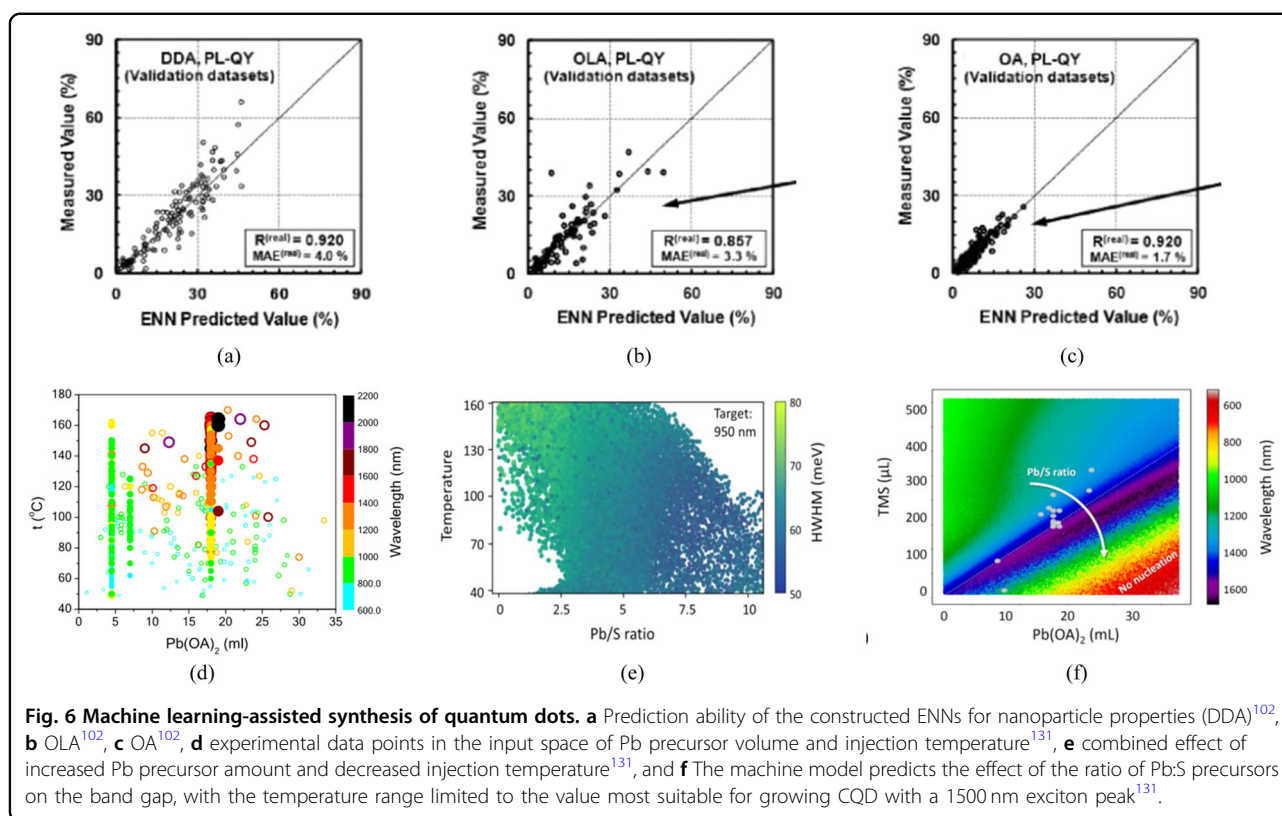
In fields such as materials science and medical imaging, processes used in synthesizing nanoparticles with specific optical properties have been found to be both expensive and time-consuming. To rapidly prepare silver nanoparticles with the desired absorption spectrum, the method of constructing predictive models by machine

Fig. 5 Machine learning-assisted synthesis of metal nanoparticles. a Algorithmic framework for a high-throughput experimental loop[124], **b** colormap of silver nanoparticles predicted by DNN framework in parameter space[124], **c** the evolution of the median fitness per generation for an extended search that achieves a single peak at 580 nm[125], **d** comparison of the spectra with the highest similarity obtained from the platform in the wavelength region of the target peak (pink)[125], and **e** TEM images of octahedral shaped gold nanorods[125].

learning was used to extract the link between chemical composition and optical properties. In this study[124], triangular nanoprisms with 50 nm edges and 10 nm heights were selected as the synthetic target, and the theoretical absorption spectrum was calculated by discrete dipole scattering (DDSCAT) plasmon resonance simulation. BO based on a Gaussian process was combined with a deep neural network (DNN) to optimize the machine learning model and improve the accuracy of the predictive model. In this study, a two-step framework-based approach was proposed to optimize nanoparticle synthesis (Fig. 5a). In the first step of the framework, the DNN regression model was trained by BO optimization. In the second step, the regression function was validated by the DNN model. The accuracy of the DNN was determined by defining a loss function. Finally, the DNN model was used to train the data gathered during the experiment and generate a spectrum for the entire parameter space. Figure 5b shows the colors generated by the DNN agent model in the plane of parameter space. For the four different regions of parameter space, the absorbance spectra predicted by the DNN were compared to experimental spectra obtained under similar conditions to indicate the predictive power of the model.

In the field of nanoparticle synthesis, the use of genetic algorithms to assist in optimizing the nanoparticles has

been widely studied. However, this is not autonomous. Salley et al.[125] proposed a robotic platform for autonomous drive syntheses of nanoparticles. The platform could be divided into multiple steps to steadily optimize the conditions used for synthesizing gold nanoparticles and explore the conditions needed for synthesizing new nanoparticle shapes while using the unique optoelectronic properties of the gold nanoparticles as the driving factor. Their synthetic method explored the chemical space for gold nanoparticle syntheses in cycles by using a genetic algorithm, an approach also known as hierarchical evolution. In the first cycle, gold nanoparticle spheres were synthesized from seeds and from the original chemical reagents. In the second cycle, gold nanoparticle rods were synthesized from seeds and proto-chemical reagents. In the third cycle, the gold nanoparticle rods synthesized in the second cycle were used as seeds to evolve and provide octahedral gold nanoparticles, thus achieving optimization of the final product. During the autonomous optimization process, a new target (the desired UV−visible spectrum with a peak maximum at 580 nm) was selected for experiments designed to explore an unknown shape mechanism. In this optimization run, 15 reactions were performed on the robotic platform until the health level remained constant. Based on iterative optimization of the genetic algorithm, comparisons of the target peak

Fig. 6 Machine learning-assisted synthesis of quantum dots. a Prediction ability of the constructed ENNs for nanoparticle properties (DDA)[102], **b** OLA[102], **c** OA[102], **d** experimental data points in the input space of Pb precursor volume and injection temperature[131], **e** combined effect of increased Pb precursor amount and decreased injection temperature[131], and **f** The machine model predicts the effect of the ratio of Pb:S precursors on the band gap, with the temperature range limited to the value most suitable for growing CQD with a 1500 nm exciton peak[131].

positions, and the resulting transmission electron microscopy images for the octahedral particles (Fig. 5c), the sample with the highest adaptability was generated from 0.479 mL of CTAB, 1.813 mL of HAuCl4, 3.327 mL of AgNO3, 3.38 mL of ascorbic acid, and 1 mL of seed solution.
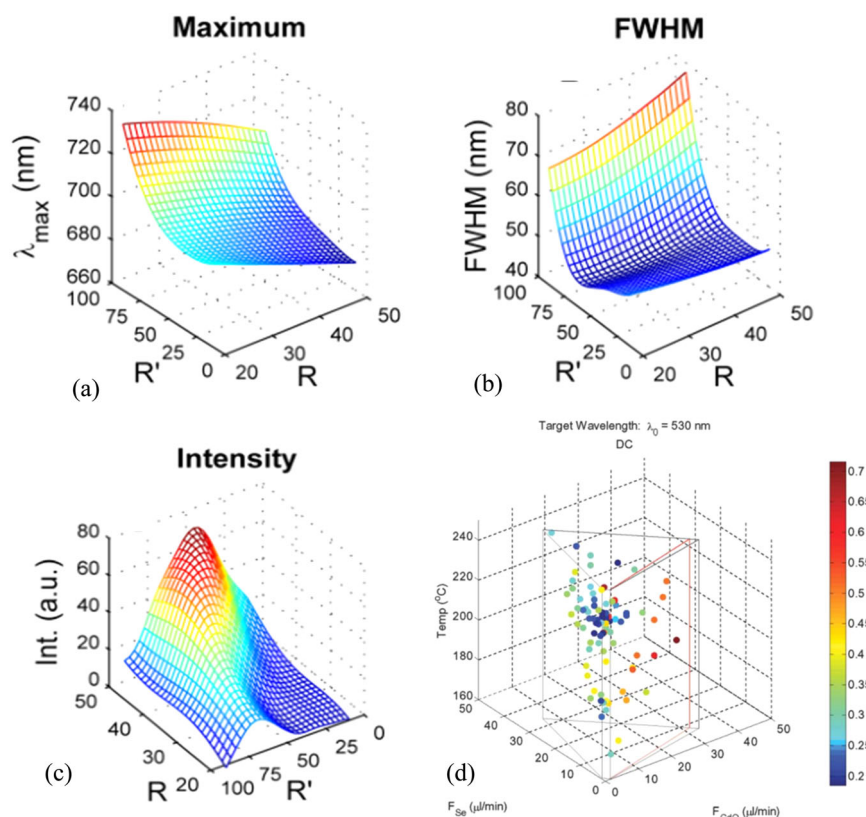
In studies of machine learning-assisted syntheses of gold nanoparticles[111], the optimal combination of synthetic conditions on a microfluidic chip can be developed autonomously and with high efficiency (fewer experiments and shorter experimental times) by using a maximum likelihood algorithm. For all experimental datasets obtained during optimization, the relationships among reaction conditions and nanoparticle properties can be constructed by machine learning.

Inorganic semiconductor nanoparticles

Inorganic semiconductor nanoparticles are being extensively studied by researchers because of their unique electronic and optical properties, such as their wide absorption ranges and narrow photoluminescence emission spectral ranges[126–128]. The unique properties of semiconductor nanoparticles enable them to play irreplaceable roles in fields such as bioimaging, solar cells, optoelectronic devices, and quantum information technology. Inorganic semiconductor nanoparticles consist of two main categories, the first of which is quantum dots.

Quantum dots can be defined as nanoparticles with dimensions close to or smaller than the Bohr exciton radius (2–10 nm) and with continuous energy bands confined to discrete energy levels (quantum confinement effect)[129,130].

Among these, colloidal quantum dots show advantages in adjusting the band gap over a wide range of visible and near-infrared spectral regions. Therefore, colloidal quantum dots play an important role in research fields such as optical sensing and photoemission. In preparing colloidal quantum dots, machine learning models can be used to construct specific combinations of parameters that achieve the desired properties[131]. The samples contained in the dataset are shown in the figure, where the lead precursor volume and injection temperature were the input variables. The solid circles in the figure are for raw data used to construct the machine learning model, and the hollow circles are for experiments performed according to the machine learning recommendations. The colors in the figure indicate the band gap wavelengths in nanometers. Among the many algorithms considered, a neural network was used to implement a BO model, build and investigate a data regression model in parameter space, and optimize the synthesis of a monodisperse product. It was found that oleylamine (OLA) affected the particle sizes and the monodispersities of colloidal quantum dots. The monodispersities of colloidal quantum

Fig. 7 Machine learning-assisted synthesis of quantum dots. a Metamodels for CdSeTe predicting luminescence maximum[132], **b** FWHM[132], **c** intensity[132], and **d** scatter plot showing the influence of F_{CdO}, F_{Se}, and the temperature on the value of the dissatisfaction coefficient for the same target wavelength of 530 nm[133].

dots could be further improved with a high Pb:S ratio and a lower injection temperature, as well as with the addition of metal chlorides. In constructing the neural network, the input layer was set up with eight inputs (outdoor temperature, $Pb(OAc)_2$ volume, OLA volume, TMS injection temperature, TMS volume, ODE volume, molar concentration of $PbCl_2$ injected at the highest temperature and molar concentration of $PbCl_2$ injected at 60 °C), two hidden layers and two outputs (band gap and linewidth). One of the machine learning models predicted the effects of Pb:S concentration ratio and temperature on the half-peak half-width and band gap, as shown in Fig. 6d–f.

During syntheses of CdSe quantum dots, the properties of the quantum dots are extremely sensitive to various reaction factors. Machine learning was used to synthesize CdSe quantum dots with precisely the desired properties. In this work[102], the ANN was trained with a total of 3404 data samples derived from six parameters for CdSe quantum dot synthetic conditions. The property relationships among the data samples were constructed with a neural network model. Three amines were used as additives: dodecylamine (DDA), OLA, and n-octylamine (OA). Neural network analyses were performed separately for

each amine dataset. Finally, an integrated neural network (ENN) was constructed by training 1600 different neural networks. The results of ENN training for the three amines are shown in Fig. 6a–c. The constructed ENN predicted the properties of the nanoparticles well even with the new experimental conditions.

CdSeTe quantum dots can be synthesized on microfluidic chips by using segmented flow capillary reactors. To construct complex relationships among the band-edge emission half height width (FWHM), maximum wavelength (maximum), and intensity (intensity) of the quantum dots and the reaction conditions, they can be optimized by fitting them with functions of the metamodeling algorithm by using kriging. The properties of the above quantum dots were taken as input to the generalized kriging algorithm, and the reaction conditions were taken as output. This allowed the output of all sample data in the prediction parameter space. The results showed that the functional relationships fitted by this algorithm had good predictive power. As shown in Fig. 7a–c, the model enabled the construction of a three-dimensional parameter space considering maximum, FWHM, and intensity for selected precursor flow ratio (R)

and residence time (Time) conditions. The model was constructed to output the optimum values of the three properties for the quantum dots. In the synthesis of CdSeTe quantum dots, the precursor flux ratios were defined by the following equations[132]:

$$R' = \frac{F_{Se}}{F_{Te} + F_{Se}} \times 100 \tag{3}$$

$$R = \frac{F_{Cd}}{F_{Cd} + F_{Te} + F_{Se}} \times 100 \tag{4}$$

where F_{Te}, F_{Cd}, and F_{Se} are the inlet flow rates of the tellurium, cadmium, and selenium precursors, respectively.

CdSe quantum dots were also prepared by injecting CdO and Se solutions into the two inlets of a heated Y-shaped microfluidic reactor. In this study, a model was designed for controlling machine learning for the synthesis of CsSe quantum dots. The data needed for machine learning were provided by monitoring the emission spectra of the obtained particles with an online spectrometer. The experimentally obtained dataset was fed into a noise-resistant global search algorithm to construct a model that iteratively updated the reaction conditions by defining a "dissatisfaction factor" (DC), which drove the model toward the desired goal and reliably output the injection rate and temperature needed to produce the optimal intensity for the selected emission wavelength. The DC is defined by the following equation[133]:

$$u_\gamma(\gamma_c) = \frac{|\gamma_c - \gamma_t|}{|\gamma_w - \gamma_t|} \tag{5}$$

where $u_\gamma(\gamma_c)$ is the dissatisfaction factor, γ_c is the current possible outcome, γ_t is the target outcome, and γ_w is the worst outcome.
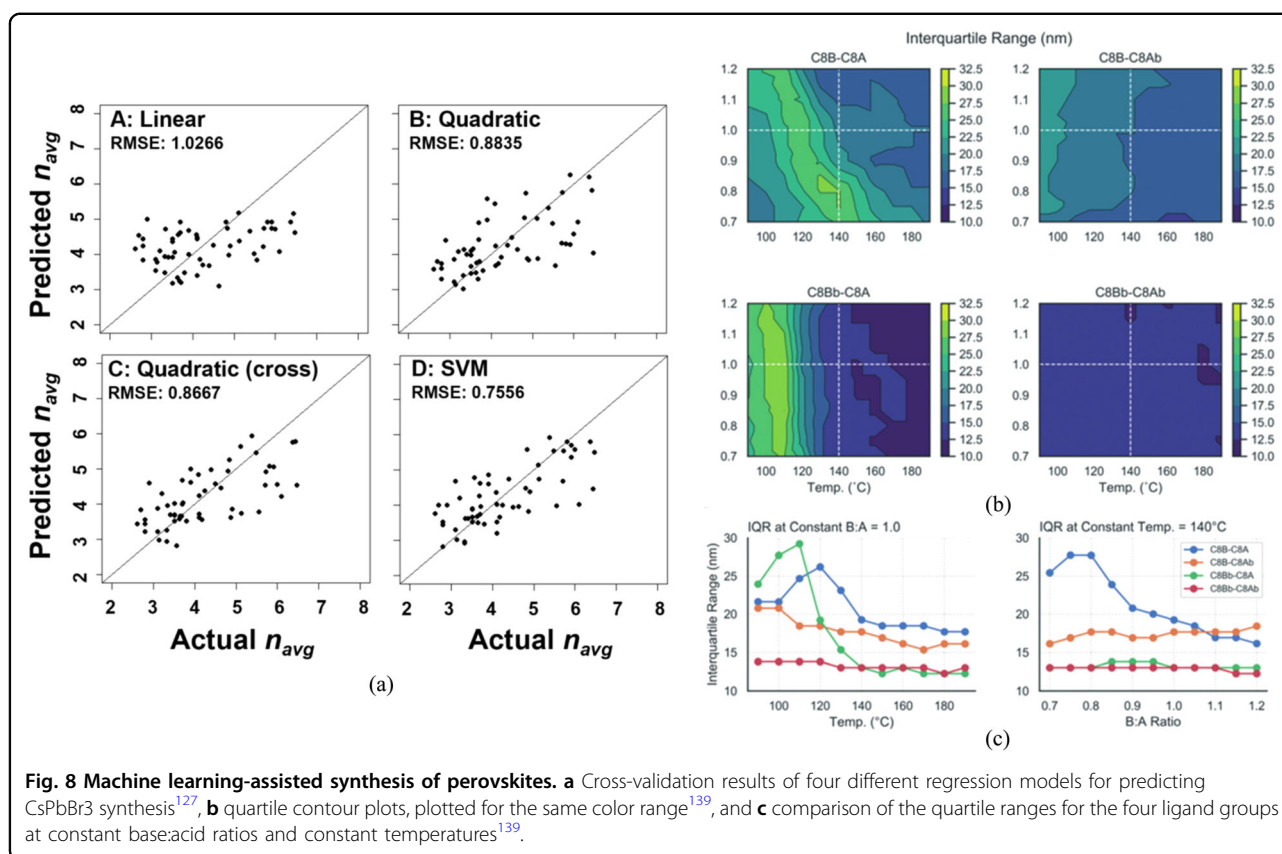
The effects of the CdO flow rate (F_{CDO}), the Se flow rate (F_{Se}), and the temperature (T) on the DC at the target wavelength of 530 nm are shown in Fig. 7d–f, in which the positions of the data points represent the corresponding values of the parameters on the three axes. The color of data points represents the value of the dissatisfaction factor. The gray line space in the figure represents the restricted ranges for flow rate and temperature. The red line represents the flow conditions that produced a 1:1 ratio of cadmium and selenium in the reaction mixture. As seen in the figure, the data points were concentrated in one place. This indicated that the algorithm preferentially sampled the low flow rate region and the high-temperature region of the parameter space. The lower DC data points in the figure are present in multiple parts of the 3D space, which indicates the presence of multiple optimal values.

Perovskites represent a second major class of inorganic semiconductor nanoparticles. With excellent optoelectronic properties and a wide range of applications in photonic devices, chalcogenide quantum dots are exciting high-priority candidate materials for rapid development and flow synthesis[134–137].

The chemical reactions used in the preparation of chalcogenide nanoparticles have numerous reaction routes and combinations that lead to significant differences in the properties of the synthesized nanoparticles. To synthesize perovskites with the desired properties, it is wise to use machine learning-assisted synthesis. For example, in the chemical reactions used for the preparation of CsPbBr3 calixarene-type nanoparticles, SVM regression models can be used to construct predictive relationships among the reaction conditions and the thickness of quantum-limited CsPbBr3 nanoplates[127]. SVM models allow for prediction, efficient dataset sample acquisition, and searches of parameter space, and they also provide fundamental insights into the roles of reaction ligands in limiting nanocrystal size. In constructing the predictive models, linear regression, quadratic regression, quadratic regression with cross terms, and SVM regression models were considered for prediction, and their predictive capabilities are shown in Fig. 8a. Machine learning-assisted syntheses of quantum dot nanoparticles provide precise localization methods for efficient navigation of the reaction design space and bring about breakthroughs in precise control of quantum dot properties and functionalities.

Optimizing the synthesis of chalcogenide nanoparticles with numerous reaction routes is one of the challenges in the field of materials science. To address this problem, a chemical reaction platform for autonomous synthesis of perovskites has been proposed[138]. This chemical reaction platform can actively synthesize inorganic chalcogenide quantum dots by using machine learning to perform the chemical synthesis efficiently and autonomously. The optimization process was as follows: first, 11 quantum dot compositions were obtained by using less than 210 mL of quantum dot solution as a starting point and without human manipulation. Second, the obtained compositions were used to pretrain machine learning with new synthetic conditions, thus accelerating the discovery of routes for successful syntheses of quantum dots. Finally, the optimization resulted in a difference between the average emission energy of the quantum dots and the target peak emission energy of less than 1 meV.

Determining optimal synthetic routes for emerging inorganic lead halide perovskites (LHPs) within their large synthetic parameter space is a challenging task. In another study[139], a modular machine learning-based microfluidic synthetic method was introduced. This method was used to intelligently navigate through the parameter space for
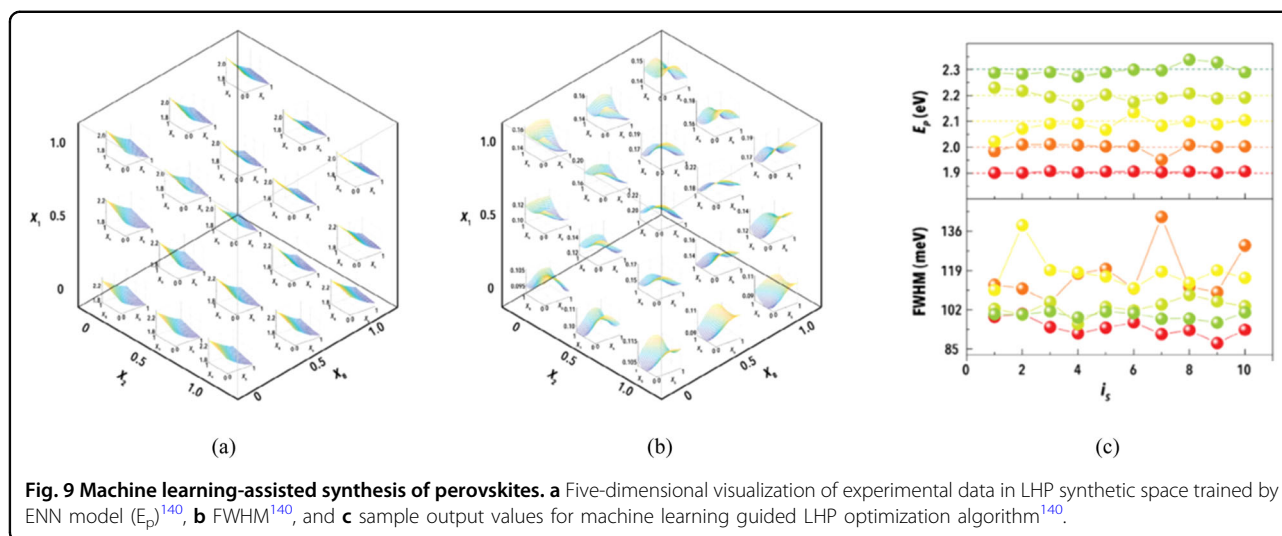
Fig. 8 Machine learning-assisted synthesis of perovskites. a Cross-validation results of four different regression models for predicting CsPbBr3 synthesis[127], **b** quartile contour plots, plotted for the same color range[139], and **c** comparison of the quartile ranges for the four ligand groups at constant base:acid ratios and constant temperatures[139].

complex LHP synthetic routes. The method used for the preparation of $CsPbBr_3$ was flow-controlled nanoparticle synthesis on a microfluidic chip with three-phase (gas−liquid−liquid) flow followed by a fast halide exchange reaction without an intermediate cleaning step. To rapidly explore the vast parameter space for synthetic routes to LHPs, a machine learning model with eight independent inputs and two output parameters was used to guide the synthesis. By using an active learning algorithm, the machine learning model determined autonomously the optimal synthetic route for CsPbX3. Prior to optimization, a five-dimensional parameter space for LHP synthesis was obtained based on the trained integrated neural network model (Fig. 8b). The exchange reaction with iodide was shown to provide the peak emission energy (Ep) and FWHM. Optimization was performed after this by using a library of archived experimental conditions developed as prophetic experience; for the target emission colors, 10 optimizations were performed by using an artificial intelligence-guided strategy, as shown in Fig. 8c. The results showed that for all targeted peak emission energy values, the AI-guided microfluidic synthetic platform was able to reach an Ep within 5 meV after only 5 experiments (40 min).

In work on the preparation of lead halide chalcogenide nanoparticles ($CsPbBr_3$), an optimization strategy based on microfluidics was used to search for the maximum nanoparticle luminescence intensity and the minimum emission bandwidth[140]. Since the luminescence intensities and the minimum emission bandwidths of nanoparticles are related to the temperature at which they are formed, the dosage ratio of reagent ligands, and the structures of ligand alkyl chains, optimization can be performed by constructing functional relationships among them. In the experiments, the ligands chosen were the linear ligands octanoic acid (C8A) and octylamine (C8B) and the split domains 2-ethylhexanoic acid (C8Ab) and 2-ethylhexylamine (C8Bb), respectively. The interquartile ranges of the four ligands at different temperatures and different ligand ratios are shown in Fig. 9, thus elucidating the effects of different ligands on reaction output variables.

Conclusion and outlook

In this paper, we introduced the characteristics and applications of nanoparticles and described the working steps and algorithms used with machine learning. Focusing mainly on metal nanoparticles and inorganic semiconductor nanoparticles, we reviewed a large number of recent studies using machine learning-assisted syntheses of functional nanoparticles. Ultimately, the analysis leads us to the following conclusions.

Fig. 9 Machine learning-assisted synthesis of perovskites. a Five-dimensional visualization of experimental data in LHP synthetic space trained by ENN model (E_p)[140], **b** FWHM[140], and **c** sample output values for machine learning guided LHP optimization algorithm[140].

(1) Syntheses of nanoparticles with expected properties require repeated and extensive experimental tests because their properties are affected by numerous reaction factors, such as the types and concentration ratios of reactants, reaction time and temperature, reaction precursors, and the structures of microchannels. Because machine learning is capable of constructing complex mapping relationships among existing data, it can be employed to assist in the syntheses of nanoparticles and solve the above problems. Machine learning models include a variety of algorithms that offer different advantages for various practical problems and the datasets that are used. In general, SVMs and ANNs are most commonly used to build predictive models. Work performed to achieve accurate predictions of nanoparticle properties with the powerful learning capability of machine learning could bring about breakthroughs in the currently limited research field.

(2) For optimization of synthetic nanoparticles and to improve the prediction accuracy of the machine learning model, an optimization algorithm can be used to optimize the hyperparameters used in the constructed model. Optimization is performed by defining a loss function such that its minimum value is the objective function. Iterative optimization by selecting suitable data based on previous experimental experience through autonomous learning can also be used.Although machine learning models can be used to synthesize nanoparticles with desired properties, there are still nonnegligible problems with the construction of datasets and with the predictive power of the algorithms. For example, the construction of large datasets is time-consuming, and the presence of incorrect or missing data in the dataset directly affects the predictive power of the machine learning model. However, in recent studies, supervised learning algorithms were used with machine learning to assist in the syntheses of nanoparticles since they are commonly used to build predictive models. Research on the use of unsupervised learning in the field of nanoparticle synthesis is relatively scarce[141-143]. With respect to the abovementioned problems, we provide an outlook with the following points:

(1) Creation and sharing of a nanoparticle database. The most important thing for machine learning before algorithm training is to prepare the dataset. The size and quality of the dataset directly affect the accuracies of the constructed models. If a comprehensive and universal database covering many nanoparticle synthetic pathways can be constructed and made openly available to researchers, it will facilitate excellent progress in the field of nanoparticle research.

(2) Research on the use of machine learning in syntheses of nanoparticles requires skill in the application of algorithms. This is very difficult for researchers who do not specialize in computer science. If the program resources of complete model algorithms can be shared on open source websites, they will provide very convenient access for researchers who do not specialize in computer applications. This would also drive the development of a further depth of content in this field of research.

(3) ANNs have powerful abilities to process images,

and if the sizes of nanoparticles can be characterized in combination with neural networks instead of by measurements with online spectrometers alone, the errors resulting from characterizing the sizes of nanoparticles can be greatly reduced.

Overall, although there are still problems with machine learning-assisted syntheses of nanoparticles, this does not detract from the fact that the technique still has great potential and offers advantages in this field.

Author contributions
X.C.: supervision, writing—review and editing, and project administration. H.L.: conceptualization, methodology, and writing—original draft.

Conflict of interest
The authors declare no competing interests.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Xie, Y. et al. Microfluidic isolation and enrichment of nanoparticles. *ACS Nano* **14**, 16220–16240 (2020).
2. Tao, H. et al. Nanoparticle synthesis assisted by machine learning. *Nat. Rev. Mater.* **6**, 701–716 (2021).
3. Paletti, P., Yue, R., Hinkle, C., Fullerton-Shirey, S. K. & Seabaugh, A. Two-dimensional electric-double-layer Esaki diode. *npj 2D Mater. Appl.* **3**, 1–7 (2019).
4. Shepherd, S. J., Issadore, D. & Mitchell, M. J. Microfluidic formulation of nanoparticles for biomedical applications. *Biomaterials* **274**, 120826 (2021).
5. Nette, J., Howes, P. D. & deMello, A. J. Microfluidic synthesis of luminescent and plasmonic nanoparticles: fast, efficient, and data-rich. *Adv. Mater. Technol.* **5**, 2000060 (2020).
6. Siavashy, S. et al. Microfluidic platform for synthesis and optimization of chitosan-coated magnetic nanoparticles in cisplatin delivery. *Carbohydr. Polym.* **265**, 118027 (2021).
7. Abedini-Nassab, R., Miandoab, M. P. & Şaşmaz, M. Microfluidic synthesis, control, and sensing of magnetic nanoparticles: a review. *Micromachines* **12**, 768 (2021).
8. Kwon, H. B., Song, W. Y., Lee, T. H., Lee, S. S. & Kim, Y. J. Monitoring the effective density of airborne nanoparticles in real time using a microfluidic nanoparticle analysis chip. *ACS Sens.* **6**, 137–147 (2021).
9. James, M., Revia, R. A., Stephen, Z. & Zhang, M. Microfluidic synthesis of iron oxide nanoparticles. *Nanomaterials* **10**, 2113 (2020).
10. Baby, T., Liu, Y., Yang, G., Chen, D. & Zhao, C. X. Microfluidic synthesis of curcumin loaded polymer nanoparticles with tunable drug loading and pH-triggered release. *J. Colloid. Interface Sci.* **594**, 474–484 (2021).
11. Yuan, N., Jin, C. & Zhang, J. X. J. Microfluidic In situ patterning of silver nanoparticles for surface-enhanced Raman spectroscopic sensing of biomolecules. *ACS Sens.* **6**, 2584–2592 (2021).
12. Nakamura, T. et al. The effect of size and charge of lipid nanoparticles prepared by microfluidic mixing on their lymph node transitivity and distribution. *Mol. Pharm.* **17**, 944–953 (2020).
13. Xu, P. F. et al. Microfluidic controllable synthesis of monodispersed sulfur nanoparticles with enhanced antibacterial activities. *Chem. Eng. J.* **398**, 125293 (2020).
14. Bachratý, H. et al. Applications of machine learning for simulations of red blood cells in microfluidic devices. *BMC Bioinforma.* **21**, 1–15 (2020).
15. Suzuki, Yuichi et al. Lipid nanoparticles loaded with ribonucleop rotein–oligonucleotide complexes synthesized using a microfluidic device exhibit robust genome editing and hepatitis B virus inhibition. *J. Control. Release* **330**, 61–71 (2021).
16. Suzuki, Y. et al. A microfluidic study to investigate the effect of magnetic iron core-carbon shell nanoparticles on displacement mechanisms of crude oil for chemical enhanced oil recovery. *J. Pet. Sci. Eng.* **184**, 106589 (2020).
17. Khizar, S. et al. Magnetic nanoparticles in microfluidic and sensing: from transport to detection. *Electrophoresis* **41**, 1206–1224 (2020).
18. Tahir, N. et al. Microfluidic fabrication and characterization of Sorafenib-loaded lipid-polymer hybrid nanoparticles for controlled drug delivery. *Int. J. Pharm.* **581**, 119275 (2020).
19. Tahir, N. et al. Synthesis of iron oxide core chitosan nanoparticles in a 3D printed microfluidic device. *J. Nanopart. Res.* **23**, 1–11 (2021).
20. Salve, M., Mandal, A., Amreen, K., Pattnaik, P. K. & Goel, S. Greenly synthesized silver nanoparticles for supercapacitor and electrochemical sensing applications in a 3D printed microfluidic platform. *Microchemical J.* **157**, 104973 (2020).
21. Zhang, K. et al. Continuous microfluidic mixing and the highly controlled nanoparticle synthesis using direct current-induced thermal buoyancy convection. *Microfluidics Nanofluidics* **24**, 1–14 (2020).
22. Lundqvist, M. et al. Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc. Natl Acad. Sci. USA* **105**, 14265–14270 (2008).
23. Li, F. et al. Multiplexed chemiluminescence determination of three acute myocardial infarction biomarkers based on microfluidic paper-based immunodevice dual amplified by multifunctionalized gold nanoparticles. *Talanta* **207**, 120346 (2020).
24. Moradikhah, F. et al. Microfluidic fabrication of alendronate-loaded chitosan nanoparticles for enhanced osteogenic differentiation of stem cells. *Life Sci.* **254**, 117768 (2020).
25. Maha, A., Srivastava, I. & Pan., D. Machine learning for precision breast cancer diagnosis and prediction of the nanoparticle cellular internalization. *ACS Sens.* **5**, 1689–1698 (2020).
26. Furxhi, I., Murphy, F., Mullins, M. & Poland, C. A. Machine learning prediction of nanoparticle in vitro toxicity: a comparative study of classifiers and ensemble-classifiers using the Copeland Index. *Toxicol. Lett.* **312**, 157–166 (2019).
27. Mohammed., A. Machine learning in protein structure prediction. *Curr. Opin. Chem. Biol.* **65**, 1–8 (2021).
28. Baichuan, S., Fernandez, M. & Barnard., A. S. Machine learning for silver nanoparticle electron transfer property prediction. *J. Chem. Inf. Modeling* **57**, 2413–2423 (2017).
29. Ryosuke, J. & Asahi, R. Predicting catalytic activity of nanoparticles by a DFT-aided machine-learning algorithm. *J. Phys. Chem. Lett.* **8**, 4279–4283 (2017).
30. Ban, Z. et al. Machine learning predicts the functional composition of the protein corona and the cellular recognition of nanoparticles. *Proc. Natl Acad. Sci. USA* **117**, 10492–10499 (2020).
31. Lievonen, M. et al. A simple process for lignin nanoparticle preparation. *Green. Chem.* **18**, 1416–1422 (2016).
32. Sahiner, N. Soft and flexible hydrogel templates of different sizes and various functionalities for metal nanoparticle preparation and their use in catalysis. *Prog. Polym. Sci.* **38**, 1329–1356 (2013).
33. Fang, F., Li, M., Zhang, J. & Lee, C. S. Different strategies for organic nano-particle preparation in biomedicine. *ACS Mater. Lett.* **2**, 531–549 (2020).
34. Chun-Long, C. & Rosi., N. L. Preparation of unique 1-D nanoparticle super-structures and tailoring their structural features. *J. Am. Chem. Soc.* **132**, 6902–6903 (2010).
35. Yalcinkaya, F., Komarek, M., Lubasova, D., Sanetrnik, F., & Maryska, J. Preparation of antibacterial nanofibre/nanoparticle covered composite yarns. *J. Nanomater.* **2016**, 7 (2016).
36. Moghaddam, K. An introduction to microbial metal nanoparticle preparation method. *J. Young Investig.* **19**, 1–7 (2010).

37. Schaeffel, D. et al. Fluorescence correlation spectroscopy directly monitors coalescence during nanoparticle preparation. *Nano Lett.* **12**, 6012–6017 (2012).

38. Swarup, R., Shankar, S. & Rhim., J.-W. Melanin-mediated synthesis of silver nanoparticle and its use for the preparation of carrageenan-based anti-bacterial films. *Food Hydrocoll.* **88**, 237–246 (2019).

39. Chang, P. R., Jian, R., Zheng, P., Yu, J. & Ma, X. Preparation and properties of glycerol plasticized-starch (GPS)/cellulose nanoparticle (CN) composites. *Carbohydr. Polym.* **79**, 301–305 (2010).

40. Cheow, W. S. & Hadinoto, K. Enhancing encapsulation efficiency of highly water-soluble antibiotic in poly (lactic-co-glycolic acid) nanoparticles: modifications of standard nanoparticle preparation methods. *Colloids Surf. A: Physicochem. Eng. Asp.* **370**, 79–86 (2010).

41. Ding, M. et al. Effect of preparation factors and storage temperature on fish oil-loaded crosslinked gelatin nanoparticle pickering emulsions in liquid forms. *Food Hydrocoll.* **95**, 326–335 (2019).

42. Ewing, C. S., Veser, G., McCarthy, J. J., Johnson, J. K. & Lambrecht, D. S. Effect of support preparation and nanoparticle size on catalyst–support interactions between Pt and amorphous silica. *J. Phys. Chem. C.* **119**, 19934–19940 (2015).

43. Li, C. et al. Preparation and biomedical applications of core–shell silica/magnetic nanoparticle composites. *J. Nanosci. Nanotechnol.* **12**, 2964–2972 (2012).

44. Chen, W., Li, S., Chen, C. & Yan, L. Self-assembly and embedding of nanoparticles by in situ reduced graphene for preparation of a 3D graphene/nanoparticle aerogel. *Adv. Mater.* **23**, 5679–5683 (2011).

45. Chiang, C.-Y., Aroh, K. & Ehrman, S. H. Copper oxide nanoparticle made by flame spray pyrolysis for photoelectrochemical water splitting–Part I. CuO nanoparticle preparation. *Int. J. Hydrog. Energy.* **37**, 4871–4879 (2012).

46. Kim, H. Y., Han, J. A., Kweon, D. K., Park, J. D. & Lim, S. T. Effect of ultrasonic treatments on nanoparticle preparation of acid-hydrolyzed waxy maize starch. *Carbohydr. Polym.* **93**, 582–588 (2013).

47. Shankar, S. & Rhim, J.-W. Preparation of sulfur nanoparticle-incorporated antimicrobial chitosan films. *Food Hydrocoll.* **82**, 116–123 (2018).

48. Ghosh, B. K., Hazra, S., Naik, B. & Ghosh, N. N. Preparation of Cu nanoparticle loaded SBA-15 and their excellent catalytic activity in reduction of variety of dyes. *Powder Technol.* **269**, 371–378 (2015).

49. Liu, S., Wang, Z., Zhang, Y., Dong, Z. & Zhang, T. Preparation of zinc oxide nanoparticle–reduced graphene oxide–gold nanoparticle hybrids for detection of NO$_2$. *RSC Adv.* **5**, 91760–91765 (2015).

50. Parolo, C. et al. Design, preparation, and evaluation of a fixed-orientation antibody/gold-nanoparticle conjugate as an immunosensing label. *ACS Appl. Mater. Interfaces.* **5**, 10753–10759 (2013).

51. Guo, Y. et al. One pot preparation of reduced graphene oxide (RGO) or Au (Ag) nanoparticle-RGO hybrids using chitosan as a reducing and stabilizing agent and their use in methanol electrooxidation. *Carbon.* **50**, 2513–2523 (2012).

52. Zhang, G. et al. Preparation of Ag-nanoparticle-loaded MnO2 nanosheets and their capacitance behavior. *Energy Fuels.* **26**, 618–623 (2012).

53. Zhang, C. et al. Preparation and tribological properties of surface-capped copper nanoparticle as a water-based lubricant additive. *Tribology Lett.* **54**, 25–33 (2014).

54. Qin, X., Lu, W., Luo, Y., Chang, G. & Sun, X. Preparation of Ag nanoparticle-decorated polypyrrole colloids and their application for H2O2 detection. *Electrochem. Commun.* **13**, 785–787 (2011).

55. Bhardwaj, H., Sumana, G. & Marquette, C. A. A label-free ultrasensitive microfluidic surface Plasmon resonance biosensor for Aflatoxin B1 detection using nanoparticles integrated gold chip. *Food Chem.* **307**, 125530 (2020).

56. Peng, B. et al. In situ surface modification of microfluidic blood–brain-barriers for improved screening of small molecules and nanoparticles. *ACS Appl. Mater. Interfaces.* **12**, 56753–56766 (2020).

57. Brzeziński, M. et al. Microfluidic-assisted nanoprecipitation of biodegradable nanoparticles composed of PTMC/PCL (co) polymers, tannic acid and doxorubicin for cancer treatment. *Colloids Surf. B: Biointerfaces.* **201**, 111598 (2021).

58. Ma, J., Yi, C. & Li, C.-W. Facile synthesis and functionalization of color-tunable Ln3+-doped KGdF4 nanoparticles on a microfluidic platform. *Mater. Sci. Eng.: C.* **108**, 110381 (2020).

59. Kimura, N. et al. Development of a microfluidic-based post-treatment process for size-controlled lipid nanoparticles and application to siRNA delivery. *ACS Appl. Mater. Interfaces.* **12**, 34011–34020 (2020).

60. Huang, F. et al. An acid-responsive microfluidic salmonella biosensor using curcumin as signal reporter and ZnO-capped mesoporous silica nanoparticles for signal amplification. *Sens. Actuators B: Chem.* **312**, 127958 (2020).

61. Zhou, J., Zhai, Y., Xu, J., Zhou, T. & Cen, L. Microfluidic preparation of PLGA composite microspheres with mesoporous silica nanoparticles for finely manipulated drug release. *Int. J. Pharm.* **593**, 120173 (2021).

62. Aghaei, H., Nazar, A. R. S. & Varshosaz, J. Double flow focusing microfluidic-assisted based preparation of methotrexate–loaded liposomal nanoparticles: encapsulation efficacy, drug release and stability. *Colloids Surf. A: Physicochem. Eng. Asp.* **614**, 126166 (2021).

63. Chang, Y. et al. Biomimetic metal-organic nanoparticles prepared with a 3D-printed microfluidic device as a novel formulation for disulfiram-based therapy against breast cancer. *Appl. Mater. Today.* **18**, 100492 (2020).

64. Omran, M., Akarri, S. & Torsaeter, O. The effect of wettability and flow rate on oil displacement using polymer-coated silica nanoparticles: a microfluidic study. *Processes.* **8**, 991 (2020).

65. Sun, H., Ren, Y., Tao, Y., Jiang, T. & Jiang, H. Three-fluid sequential micromixing-assisted nanoparticle synthesis utilizing alternating current electrothermal flow. *Ind. Eng. Chem. Res.* **59**, 12514–12524 (2020).

66. Cao, Y. et al. Microfluidic manufacturing of SN-38-loaded polymer nanoparticles with shear processing control of drug delivery properties. *Mol. Pharm.* **16**, 96–107 (2018).

67. Kašpar, O., Koyuncu, A. H., Pittermannová, A., Ulbrich, P. & Tokárová, V. Governing factors for preparation of silver nanoparticles using droplet-based microfluidic device. *Biomed. Microdevices.* **21**, 1–14 (2019).

68. Kumar, D. V., Ravi, B. L. V., Prasad & Kulkarni, A. A. Segmented flow synthesis of Ag nanoparticles in spiral microreactor: role of continuous and dispersed phase. *Chem. Eng. J.* **192**, 357–368 (2012).

69. Solomun, J. I., Totten, J. D., Wongpinyochit, T., Florence, A. J. & Seib, F. P. Manual versus microfluidic-assisted nanoparticle manufacture: impact of silk fibroin stock on nanoparticle characteristics. *ACS Biomater. Sci. Eng.* **6**, 2796–2804 (2020).

70. Riewe, J. et al. Antisolvent precipitation of lipid nanoparticles in microfluidic systems–a comparative study. *Int. J. Pharm.* **579**, 119167 (2020).

71. Wang, X., Liu, L., Zhang, W. & Ma, X. Prediction of plant uptake and translocation of engineered metallic nanoparticles by machine learning. *Environ. Sci. Technol.* **55**, 7491–7500 (2021).

72. Wang, M., Wang, T., Cai, P. & Chen, X. Nanomaterials discovery and design through machine learning. *Small Methods.* **3**, 1900025 (2019).

73. Yu, F., Wei, C., Deng, P., Peng, T. & Hu, X. Deep exploration of random forest model boosts the interpretability of machine learning studies of complicated immune responses and lung burden of nanoparticles. *Sci. Adv.* **7**, eabf4130 (2021).

74. Egorov, E., Pieters, C., Korach-Rechtman, H., Shklover, J. & Schroeder, A. Robotics, microfluidics, nanotechnology and AI in the synthesis and evaluation of liposomes and polymeric drug delivery systems. *Drug Deliv. Transl. Res.* **11**, 345–352 (2021).

75. Juganson, K., Ivask, A., Blinova, I., Mortimer, M. & Kahru, A. NanoE-Tox: new and in-depth database concerning ecotoxicity of nanomaterials. *Beilstein J. Nanotechnol.* **6**, 1788–1804 (2015).

76. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A. & Kim, C. Machine learning in materials informatics: recent applications and prospects. *npj Comput. Mater.* **3**, 1–13 (2017).

77. Ban, Z. et al. Machine learning predicts the functional composition of the protein corona and the cellular recognition of nanoparticles. *Proc. Natl Acad. Sci. USA.* **117**, 10492–10499 (2020).

78. Barnard, A. S. & Opletal, G. Predicting structure/property relationships in multi-dimensional nanoparticle data using t-distributed stochastic neighbour embedding and machine learning. *Nanoscale.* **11**, 23165–23172 (2019).

79. Gasper, R., Shi, H. & Ramasubramaniam, A. Adsorption of CO on low-energy, low-symmetry Pt nanoparticles: energy decomposition analysis and prediction via machine-learning models. *J. Phys. Chem. C.* **121**, 5612–5619 (2017).

80. Kingston, B. R., Syed, A. M., Ngai, J., Sindhwani, S. & Chan, W. C. Assessing micrometastases as a target for nanoparticles using 3D microscopy and machine learning. *Proc. Natl Acad. Sci. USA.* **116**, 14937–14946 (2019).

81. Zheng, B. & Gu, G. X. Machine learning-based detection of graphene defects with atomic precision. *Nano-micro Lett.* **12**, 1–13 (2020).

82. Srikanth, S., Dubey, S. K., Javed, A. & Goel, S. Droplet based microfluidics integrated with machine learning. *Sens. Actuators A: Phys.* **332**, 113096 (2021).

83. Joshi, K. et al. A machine learning-assisted nanoparticle-printed biochip for real-time single cancer cell analysis. *Adv. Biosyst.* **4**, 2000160 (2020).

84. Lu, W., Xiao, R., Yang, J., Li, H. & Zhang, W. Data mining-aided materials discovery and optimization. *J. Materiomics* **3**, 191–201 (2017).

85. Volk, A. A., Robert, W. E. & Abolhasani, M. Accelerated development of colloidal nanomaterials enabled by modular microfluidic reactors: toward autonomous robotic experimentation. *Adv. Mater.* **33**, 2004495 (2021).

86. Sarkar, S. et al. Machine learning-aided quantification of antibody-based cancer immunotherapy by natural killer cells in microfluidic droplets. *Lab a Chip* **20**, 2317–2327 (2020).

87. Lashkaripour, A. et al. Machine learning enables design automation of microfluidic flow-focusing droplet generation. *Nat. Commun.* **12**, 1–14. (2021).

88. Tao, Q., Xu, P., Li, M. & Lu, W. Machine learning for perovskite materials design and discovery. *npj Comput. Mater.* **7**, 1–18 (2021).

89. Lazarovits, J. et al. Supervised learning and mass spectrometry predicts the in vivo fate of nanomaterials. *ACS Nano* **13**, 8023–8034 (2019).

90. Nanba, Y. & Koyama, M. NO adsorption on 4d and 5d transition-metal (Rh, Pd, Ag, Ir, and Pt) nanoparticles: density functional theory study and supervised learning. *J. Phys. Chem. C.* **123**, 28114–28122 (2019).

91. Shin, Y. et al. Microfluidic multi-scale homogeneous mixing with uniform residence time distribution for rapid production of various metal core–shell nanoparticles. *Adv. Funct. Mater.* **31**, 2007856 (2021).

92. Zuranski, A. M., Martinez Alvarado, J. I., Shields, B. J. & Doyle, A. G. Predicting reaction yields via supervised learning. *Acc. Chem. Res.* **54**, 1856–1865 (2021).

93. Ouassil, N., Pinals, R. L., Del Bonis-O'Donnell, J. T., Wang, J. W. & Landry, M. P. Supervised learning model predicts protein adsorption to carbon nanotubes. *Sci. Adv.* **8**, eabm0898 (2022).

94. Timoshenko, J. et al. Probing atomic distributions in mono-and bimetallic nanoparticles by supervised machine learning. *Nano Lett.* **19**, 520–529 (2018).

95. Zubair, G. et al. Intelligent supervised learning for viscous fluid submerged in water based carbon nanotubes with irreversibility concept. *Int. Commun. Heat. Mass Transf.* **130**, 105790 (2022).

96. Jahed Armaghani, D. et al. Examining hybrid and single SVM models with different kernels to predict rock brittleness. *Sustainability* **12**, 2229 (2020).

97. Dies, H., Raveendran, J., Escobedo, C. & Docoslis, A. Rapid identification and quantification of illicit drugs on nanodendritic surface-enhanced Raman scattering substrates. *Sens. Actuators B: Chem.* **257**, 382–388 (2018).

98. Xing, W. & Bei, Y. Medical health big data classification based on KNN classification algorithm. *IEEE Access* **8**, 28808–28819 (2019).

99. Khorrami, Gh. H., Kompany, A. & Khorsand Zak, A. A facile sol–gel approach to synthesize KNN nanoparticles at low temperature. *Mater. Lett.* **110**, 172–175 (2013).

100. Jumin, E., Basaruddin, F. B., Yusoff, Y. B., Latif, S. D. & Ahmed, A. N. Solar radiation prediction using boosted decision tree regression model: a case study in Malaysia. *Environ. Sci. Pollut. Res.* **28**, 26571–26583 (2021).

101. Le, B. H. & Seo, Y. J. Highly sensitive MicroRNA 146a detection using a gold nanoparticle–based CTG repeat probing system and isothermal amplification. *Analytica Chim. Acta* **999**, 155–160 (2018).

102. Zhang, Q., Yu, H., Barbiero, M., Wang, B. & Gu, M. Artificial neural networks enabled by nanophotonics. *Light.: Sci. Appl.* **8**, 1–14 (2019).

103. Riordon, J., Sovilj, D., Sanner, S., Sinton, D. & Young, E. W. Deep learning with microfluidics for biotechnology. *Trends Biotechnol.* **37**, 310–324 (2019).

104. Liu, R. et al. Causal inference machine learning leads original experimental discovery in CdSe/CdS core/shell nanoparticles. *J. Phys. Chem. Lett.* **11**, 7232–7238 (2020).

105. Xiong, J. et al. Enhancing privacy and availability for data clustering in intelligent electrical service of IoT. *IEEE Internet Things J.* **6**, 1530–1540 (2018).

106. Sizochenko, N., Syzochenko, M., Fjodorova, N., Rasulev, B. & Leszczynski, J. Evaluating genotoxicity of metal oxide nanoparticles: application of advanced supervised and unsupervised machine learning techniques. *Ecotoxicol. Environ. Saf.* **185**, 109733 (2019).

107. Parker, A. J. & Amanda, S. B. Machine learning reveals multiple classes of diamond nanoparticles. *Nanoscale Horiz.* **5**, 1394–1399 (2020).

108. Wang, Z. et al. Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* **1**, 137–145 (2018).

109. Artrith, N., Urban, A. & Ceder, G. Constructing first-principles phase diagrams of amorphous Li x Si using machine-learning-assisted sampling with an evolutionary algorithm. *J. Chem. Phys.* **148**, 241711 (2018).

110. Bawazer, L. A. et al. Combinatorial microfluidic droplet engineering for biomimetic material synthesis. *Sci. Adv.* **2**, e1600567 (2016).

111. Tao, H. et al. Self-driving platform for metal nanoparticle synthesis: combining microfluidics and machine learning. *Adv. Funct. Mater.* **31**, 2106725 (2021).

112. Niculescu, A. G., Chircov, C., Bîrcă, A. C. & Grumezescu, A. M. Fabrication and applications of microfluidic devices: a review. *Int. J. Mol. Sci.* **22**, 2011 (2021).

113. Cheng, Y., Da Ling, S., Geng, Y., Wang, Y. & Xu, J. Microfluidic synthesis of quantum dots and their applications in bio-sensing and bio-imaging. *Nanoscale Adv.* **3**, 2180–2195 (2021).

114. Maceiczyk, R. M. et al. Microfluidic reactors provide preparative and mechanistic insights into the synthesis of formamidinium lead halide perovskite nanocrystals. *Chem. Mater.* **29**, 8433–8439 (2017).

115. Yang, S. Y., Cheng, F. Y., Yeh, C. S. & Lee, G. B. Size-controlled synthesis of gold nanoparticles using a micro-mixing system. *Microfluidics Nanofluidics* **8**, 303–331 (2010).

116. Saha, N., Astray, G. & Gupta, S. D. Modelling and optimization of biogenic synthesis of gold nanoparticles from leaf extract of Swertia chirata using artificial neural network. *J. Clust. Sci.* **29**, 1151–1159 (2018).

117. Aldakov, D. & Reiss, P. Safer-by-design fluorescent nanocrystals: Metal halide perovskites vs semiconductor quantum dots. *J. Phys. Chem. C.* **123**, 12527–12541 (2019).

118. Lin, L. et al. Microfluidic fabrication of fluorescent nanomaterials: a review. *Chem. Eng. J.* **425**, 131511 (2021).

119. Zhou, Z., Li, X. & Zare, R. N. Optimizing chemical reactions with deep reinforcement learning. *ACS Cent. Sci.* **3**, 1337–1344 (2017).

120. Dressler, O. J., Howes, P. D., Choo, J. & deMello, A. J. Reinforcement learning for dynamic microfluidic control. *ACS Omega* **3**, 10084–10091 (2018).

121. Gherman, A. M. M. et al. Artificial neural networks modeling of the parameterized gold nanoparticles generation through photo-induced process. *Mater. Res. Express* **5**, 085011 (2018).

122. Mirzaei, M., Furxhi, I., Murphy, F. & Mullins, M. A machine learning tool to predict the antibacterial capacity of nanoparticles. *Nanomaterials* **11**, 1774 (2021).

123. Sugano, K. et al. Mixing speed-controlled gold nanoparticle synthesis with pulsed mixing microfluidic system. *Microfluidics Nanofluidics* **9**, 1165–1174 (2010).

124. Mekki-Berrada, F. et al. Two-step machine learning enables optimized nanoparticle synthesis. *npj Comput. Mater.* **7**, 1–10 (2021).

125. Salley, D. et al. A nanomaterials discovery robot for the Darwinian evolution of shape programmable gold nanoparticles. *Nat. Commun.* **11**, 1–7 (2020).

126. Lignos, I., Maceiczyk, R. & deMello., A. J. Microfluidic technology: uncovering the mechanisms of nanocrystal nucleation and growth. *Acc. Chem. Res.* **50**, 1248–1257 (2017).

127. Braham, E. J. et al. Machine learning-directed navigation of synthetic design space: a statistical learning approach to controlling the synthesis of perovskite halide nanoplatelets in the quantum-confined regime. *Chem. Mater.* **31**, 3281–3292 (2019).

128. Peng, J., Muhammad, R., Wang, S. L. & Zhong, H. Z. How machine learning accelerates the development of quantum dots? *Chin. J. Chem.* **39**, 181–188 (2021).

129. Regonia, P. R. et al. Predicting the band gap of ZnO quantum dots via supervised machine learning models. *Optik* **207**, 164469 (2020).

130. Tang, B. et al. Machine learning-guided synthesis of advanced inorganic materials. *Mater. Today* **41**, 72–80 (2020).

131. Voznyy, O. et al. Machine learning accelerates discovery of optimal colloidal quantum dot synthesis. *ACS Nano* **13**, 11122–11128 (2019).

132. Maceiczyk, R. M., Andrew & deMello., J. Fast and reliable metamodeling of complex reaction spaces using universal kriging. *J. Phys. Chem. C.* **118**, 20026–20033 (2014).

133. Krishnadasan, S., Brown, R. J. C., Demello, A. J. & Demello, J. C. Intelligent routes to the controlled synthesis of nanoparticles. *Lab a Chip* **7**, 1434–1441 (2007).

134. Sun, S. et al. Accelerated development of perovskite-inspired materials via high-throughput synthesis and machine-learning diagnosis. *Joule* **3**, 1437–1451 (2019).

135. Higgins, K., Valleti, S. M., Ziatdinov, M., Kalinin, S. V. & Ahmadi, M. Chemical robotics enabled exploration of stability in multicomponent lead halide perovskites via machine learning. *ACS Energy Lett.* **5**, 3426–3436 (2020).

136. Zhang, L., He, M. & Shao, S. Machine learning for halide perovskite materials. *Nano Energy* **78**, 105380 (2020).

137. Balachandran, P. V., Kowalski, B., Sehirlioglu, A. & Lookman, T. Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning. *Nat. Commun.* **9**, 1–9 (2018).

138. Epps, R. W. et al. Artificial chemist: an autonomous quantum dot synthesis bot. *Adv. Mater.* **32**, 2001626 (2020).

139. Li, S. et al. Automated microfluidic screening of ligand interactions during the synthesis of cesium lead bromide nanocrystals. *Mol. Syst. Des. Eng.* **5**, 1118–1130 (2020).

140. Abdel-Latif, K. et al. Self-driven multistep quantum dot synthesis enabled by autonomous robotic experimentation in flow. *Adv. Intell. Syst.* **3**, 2000245 (2021).

141. Kolenov, D., Davidse, D., Le Cam, J. & Pereira, S. F. Convolutional neural network applied for nanoparticle classification using coherent scatterometry data. *Appl. Opt.* **59**, 8426–8433 (2020).

142. Shalaby, K. S. et al. Determination of factors controlling the particle size and entrapment efficiency of noscapine in PEG/PLA nanoparticles using artificial neural networks. *Int. J. Nanomed.* **9**, 4953 (2014).

143. Russo, D. P. et al. Virtual molecular projections and convolutional neural networks for the end-to-end modeling of nanoparticle activities and properties. *Anal. Chem.* **92**, 13971–13979 (2020).