Prediction and Regression Analysis of MSW by Using Hypothesis Test Method

Himanshu Jha[1], R. K. Bhatia[2]

[1]*ME Scholar of Environment Engineering, Department of Civil Engineering, JEC Jabalpur*

[2]*Associate Professor, Department of Civil Engineering, JEC Jabalpur*

Jabalpur Engineering College (JEC) Jabalpur, Department of Civil Engineering. MP, India

---***---

Abstract- Population growth and the acceleration of urbanization have led to a sharp increase in municipal solid waste production, and researchers have sought to use advanced technology to solve this problem. Machine learning (ML) algorithms are good at modeling complex nonlinear processes and have been gradually adopted to promote municipal solid waste management (MSWM) and help the sustainable development of the environment in the past few years. In this study, more than 200 publications published over the last two decades (1999-2016) were reviewed and analyzed. This paper summarizes the application of ML algorithms in the whole process of MSWM, from waste generation to collection and transportation, to final disposal. Through this comprehensive review, the gaps and future directions of ML application in MSWM are discussed, providing theoretical and practical guidance for follow-up related research.

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

1. INTRODUCTION

There are a few issues that are given less attention and solid waste management (SWM) is one of those municipalities' programs. In every region of India, the amount of municipal solid waste generated is increasing rapidly due to population growth and urbanization. This issue is presenting many challenges to authorities in solid waste management. Due to limited financial resources and other difficulties, this issue needed serious attention and hence the Municipal Solid Waste Rules, 2000. It is applied to every municipal authority in India and these rules were mandatory for the administrative authorities to undertake responsibilities for all activities relating to municipal solid waste management. Solid waste management activities can create new jobs and help in boosting the economy. There are many use-cases of this solid waste like it can be used to produce electricity while mitigating the environmental impacts.

A survey of MSWM practices in Indian municipality corporations suggests that the underestimation of generation rates is a major concern in India and therefore, other problems like lack of use of technology, lack of reliable information, and underestimating resource requirements. Before the corona outbreak in the year 2020, the Indian economy was growing at a good rate. India is a developing country and there is a vast history of developing countries that are likely to help in growing the rates of solid waste generation.

2. RESEARCH MOTIVATION

The present approaches for MSW management are resulting in inefficient utilization of time and resources because of the uncertainty in solid waste generation. There are a lot of modern technologies and tools like digitization, planting sensors, and solid waste optimized disposal systems that can be used to do efficient waste management mainly towards optimal utilization of time and resources. But all such waste management efforts are facing issues because of many challenges. All this starts with the incapability of local authorities handling solid waste in a better way. The main obstacle to this is that waste management is done at disaggregate levels and it involves different channels, making the data collection and compilation difficult. Many factors are contributing to this like a rapidly increasing population, building new factories, shops, and restaurants, frequent public gatherings, etc. Hence judging the amount of waste generation has become a very difficult task resulting in inefficient waste management.

3. BACKGROUND OF MACHINE LEARNING METHODS

Machine learning is the core area of Artificial Intelligence (Dasgupta &Nath, 2016). Machine learning and artificial intelligence has merged with big data analytics has improved performance computing to create new opportunities for data intensive science in the multi-disciplinary domain (Liakoset al., 2018).

It employs variety of statistical, probabilistic and optimization techniques that allows computers to "learn" from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets (Cruz & Wishart, 2006).

In this paper, we present a comprehensive review of the application of ML in solid waste management (SWM). A number of relevant papers are presented that emphasize key and unique features of popular ML models. The paper is structured in sections as follows: Section 2 the definition of terms used in ML, most common models and algorithms. Section 3 presents the implemented methodology for the collection and categorization of the presented works. Finally, in Section 4, the strengths derived from the implementation of ML in SWM are listed, as well as the future expectations in the domain. Machine learning is a branch of artificial intelligence research that employs a variety of statistical, probabilistic and optimization tools to "learn" from past examples and to then use that prior training to classify new data, identify new patterns or predict novel trends (Mitchell, 1997).

According to (Liakoset al., 2018), ML methodologies involve a learning process with the objective to learn from "experience" (training data)to perform a task. Machine learning, like statistics, is used to analyze and interpret data. Unlike statistics, though, machine learning methods can employ Boolean logic, absolute conditionality, conditional probabilities and unconventional optimization strategies to model data or classify patterns (Cruz, & Wishart, 2006). Data in ML consists of a set of examples. Usually, an individual example is described by a set of attributes, also known as features or variables. A feature can be nominal (enumeration), binary, ordinal, or numeric.

4. OVERVIEW OF ML

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

(2) Choosing the proper ML algorithms and parameters, and

(3) Testing and evaluating performance.

5. PROBLEM FORMULATION AND METHOD SELECTION CRITERIA

In summary, the ANN models could be applied for modeling MSW quantities given limited data. The flexibility of neural network tools is a feature that allows for the consideration of a variety of additional factors, including economic, demographic, geographic, social, technological, legislative, and administrative, all of which may play a role in determining the final quantity of municipal wastes. Despite the excellent ability of standalone ANN models, their training algorithms may trap in local optimums or may be slow to convergence. Optimization algorithms are regarded as viable alternatives to standard training algorithms because they avoid trapping in local optimums [1,2]. However, limited research was undertaken to examine the performance of hybrid ANN models to forecast MSW generation quantities. Furthermore, it is necessary to assess the performances of pure and hybrid ANN models using numerous evaluation indices. In this regard, this research involves training and evaluating the performance of conventional and hybrid ANN models in predicting MSW quantities. Because of its high prediction accuracy and consistency over time, the ANN is coupled with the particle swarm optimization (PSO) algorithm in this research study. The capacity to estimate MSW at a city level allows local waste management organizations and government agencies to develop robust waste management strategies. In addition, the developed prediction model may be incorporated into legal laws and systems to reduce municipal waste generation rates, improve the efficiency of waste energy recovery, and improve adherence to sustainable development principles. The major contributions of this research could be summarized as follows:

1. Incorporating the influence of economic, demographic, and social aspects on the quantity of generated wastes.

2. Estimating the waste quantities using traditional and hybrid neural network models and comparing their performances using several evaluation metrics.

3. Enhancing the performance metrics of the developed prediction models in the literature.

OBJECTIVES

Following objectives are perform given research work-

1. To collection of previous year data.

2. To identification of ANN results.

3. To improvement of ANN prediction accuracy.

6. PROPOSED METHODOLOGY

ANNs have been used to accelerate reliability analysis of infrastructures subject to natural disasters and to predict foundation settlements. ANNs have also been used for building black-box models in geoscience: hydrology, ocean modelling and coastal engineering, and geomorphology. ANNs have been

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

employed in cybersecurity, with the objective to discriminate between legitimate activities and malicious ones. For example, machine learning has been used for classifying Android malware,[3] for identifying domains belonging to threat actors and for detecting URLs posing a security risk.[4] Research is underway on ANN systems designed for penetration testing, for detecting botnets,[4] credit cards frauds[6] and network intrusions.

ANNs have been proposed as a tool to solve partial differential equations in physics[7][5][9] and simulate the properties of many-body open quantum systems.In brain research ANNs have studied short-term behavior of individual neurons,[2] the dynamics of neural circuitry arise from interactions between individual neurons and how behavior can arise from abstract neural modules that represent complete subsystems. Studies considered long-and short-term plasticity of neural systems and their relation to learning and memory from the individual neuron to the system level.

One type of network sees the nodes as ‘artificial neurons’. These are called artificial neural networks (ANNs). An artificial neuron is a computational model inspired in the natural neurons. Natural neuron signals through synapses located on the dendrites or membrane of the neuron. When the signals received are strong enough (surpass a certain threshold), the neuron is activated and emits a signal though the axon. This signal might be sent to another synapse, and might activate other neurons.

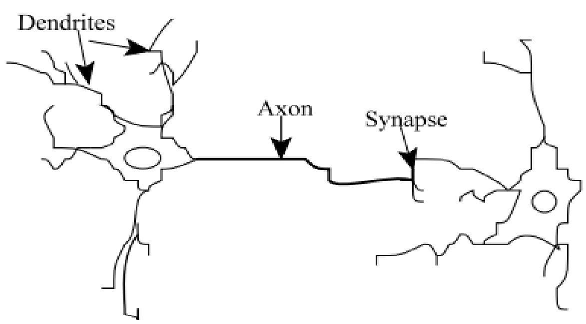

Fig -1: **Natural neurons (artist’s conception).**

The complexity of real neurons is highly abstracted when modelling artificial neurons. These basically consist of inputs (like synapses), which are multiplied by weights (strength of the respective signals), and then computed by a mathematical function which determines the activation of the neuron. Another function (which may be the identity) computes the output of the artificial neuron (sometimes in dependance of a certain threshold). ANNs combine artificial neurons in order to process information.

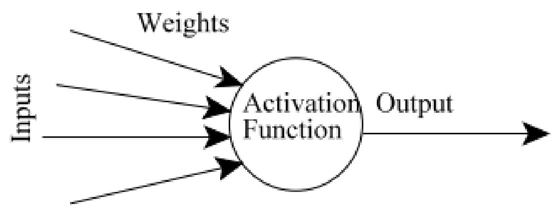

Fig -2: **An artificial neuron.**

The higher a weight of an artificial neuron is, the stronger the input which is multiplied by it will be. Weights can also be negative, so we can say that the signal is inhibited by the negative weight. Depending on the weights, the computation of the neuron will be different. By adjusting the weights of an artificial neuron we can obtain the output we want for specific inputs. But when we have an ANN of hundreds or thousands of neurons, it would be quite complicated to find by hand all the necessary weights. But we can find algorithms which can adjust the weights of the ANN in order to obtain the desired output from the network. This process of adjusting the weights is called learning or training. The number of types of ANNs and their uses is very high. Since the first neural model by McCulloch and Pitts (1943) there have been developed hundreds of different models considered as ANNs. The differences in them might be the functions, the accepted values, the topology, the learning algorithms, etc. Also there are many hybrid models where each neuron has more properties than the ones we are reviewing here. Because of matters of space, we will present only an ANN which learns using the backpropagation algorithm (Rumelhart and McClelland, 1986) for learning the appropriate weights, since it is one of the most common models used in ANNs, and many others are based on it. Since the function of ANNs is to process information, they are used mainly in fields related with it. There are a wide variety of ANNs that are used to model real neural networks, and study behaviour and control in animals and machines, but also there are ANNs which are used for engineering purposes, such as pattern recognition, forecasting, and data compression[10].

Neural Network Algorithms – Artificial Neural Networks arguably works close enough to the human brain. Conceptually artificial neural networks are inspired by neural networks in the brain but the actual implementation in machine learning is way far from reality. ANN take in multiple inputs and produce a single output. Point to note ANN’s are inspired by the animal brain, but nowhere close to biological neural networks [11]. Neural Network Work Flow – Layers of Learning

Neural networks learning process is not very different from humans, humans learn from experience in lives while neural networks require data to gain experience and learn. Accuracy increases with the amount of data over time. Similarly, humans also perform the same task better and better by doing any task you do over and over [12]. Neural Network Algorithms' underlying foundation of neural networks is a layer and layers of connections. The entire neural network model is based on a layered architecture. Each layer has its own responsibility. These networks are designed to make use of layers of "neurons" to process raw data, find patterns into it and objects which are usually hidden to naked eyes. To train a neural network, data scientist put their data in three different baskets [13].

Training data set – This helps networks to understand and know the various weights between nodes.

Validation data set – To fine-tune the data sets.

Test data set – To evaluate the accuracy and records margin of error.

Layer takes input, extract feature and feed into the next layer i.e. each layer work as an input layer to another layer. This is to receive information and last layer job is to throw output of the required information. Hidden layers or core layers process all the information in between.

• Assign a random weight to all the links to start the algorithm [2].

• Find links the activation rate of all hidden nodes by using the input and links.

• Find the activation rate of output nodes with the activation rate of hidden nodes and link to output.

• Errors are discovered at the output node and to recalibrate all the links between hidden & output nodes.

• Using the weights and error at the output; cascade down errors to hidden & output nodes. Weights get applied on connections as the best friend for neural networks.

• Recalibrate & repeat the process of weights between hidden and input nodes until the convergence criteria are met.

• Finally the output value of the predicted value or the sum of the three output values of each neuron. This is the output.

• Patterns of information are fed into the network via the input units, which trigger the layers of hidden units, and these, in turn, arrive at the output units [14].

The Backpropagation Algorithm

The backpropagation algorithm (Rumelhart and McClelland, 1986) is used in layered feed-forward ANNs. This means that the artificial neurons are organized in layers, and send their signals "forward", and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers. The backpropagation algorithm uses supervised learning, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea of the backpropagation algorithm is to reduce this error, until the ANN learns the training data. The training begins with random weights, and the goal is to adjust them so that the error will be minimal.

7. DATA PRE-PROCESSING AND FEATURE SELECTION

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

In this study, machine learning approaches (ANN) were applied with training and test datasets regarding TOTAL WASTE collection per year inputs/parameters. The algorithms were tested with datasets, which belong to 5 years (1999-2026), and the cites mostly INDIA. The results show that ANN methods can be applied successfully to the City in INDIA as well as for the case of INDIA to predict the MSW generation.

Following parameters are evaluating by-

1. **MSE**
2. **R^2 Values**

8. RESULT AND SIMULATION

The ANN method were implemented by using the appropriate MATLAB 2015 functions of feedforwardnet(), newgrnn(), and fitrsvm(), respectively. The running times of the algorithm can be considered almost the same in the training stage for this problem. **Result** mentioned the performance of the ANN with the training datasets (2015-2020) by analysis of the predicted MSW values by the models to the known actual MSW accuracy values. Following points are obtained results-

Table -1: DATA COLLECTION TABLE

City/year	1999-2000	2004-05	2010-11	2015-16
Mumbai (Mh)	5355	5320	6500	11000
Delhi i	400	5922	6800	8700
Bangalore (Krn)	200	1669	3700	3700
Chennai (TN)	3124	3036	4500	5000
Hyderabad (Tel)	1566	2187	4200	4000
Ahmedabad (Guj)	1683	1302	2300	2500

Kolkata (WB)	3692	2653	3670	4000
Surat (Guj)	900	1000	1200	1680
Pune(Mah)	700	1175	1300	1600
Jaipur (Raj)	580	904	310	1000
Luck now (UP)	1010	475	1200	1200
Kanpur (UP)	1200	1100	1600	1500
Nagpur (Mh)	443	504	650	1000
Visakhapatnam (AP)	300	584	334	350
Indore (MP)	350	557	720	850
Bhopal (MP)	546	574	350	700
Patna (Bhr)	330	511	220	450
Vadodara (Guj)	400	357	600	700
Ludhiana (Pb)	400	735	850	850
Coimbatore (TN)	350	530	700	850
Agra (UP)	-	654	520	790
Madurai (TN)	370	275	450	450
Nashik (Mh)	-	200	350	500
Vijayawada (AP)	-	374	600	550
Faridabad (Hr)	-	448	700	400
Meerut (UP)	-	490	520	500
Rajkot (Guj)	-	207	230	450
Kalian-dombivali	-	-	510	650
Varanasi (UP)	412	425	450	500
Srinagar (JK)	-	428	550	550
Dhanbad (Jh)	-	77	150	180
Amritsar (Pb)	-	438	550	600
Allahabad (UP)	-	509	350	450
Ranchi (Jh)	-	208	140	150
Jabalpur (MP)	148	216	400	550
Gwalior (MP)	-	178	285	300
Raipur (Chh)	-	184	224	230

3. MADHYA PRADESH CITY

Table -2: MADHYA PRADESH CITY.

City/year	1999-2000	2004-05	2010-11	2015-16
Indore (MP)	350	557	720	850
Bhopal (MP)	546	574	350	700
Jabalpur (MP)	148	216	400	550
Gwalior (MP)	10	178	285	300
Dewas (MP)	258	136	235	436

4. Indore (MP) 2025 Prediction Session ANN results

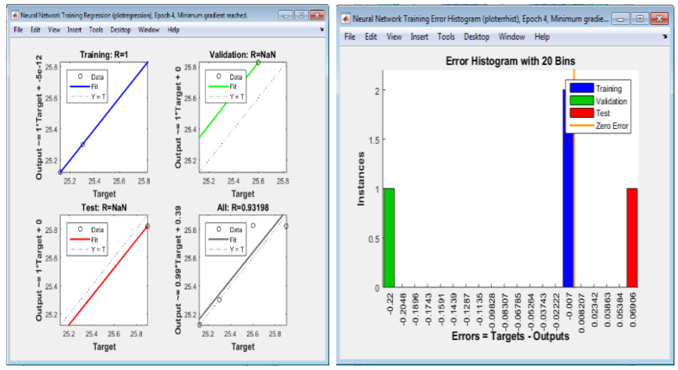

Fig -3: Indore (MP) 2025 Prediction Session ANN results.

5. Indore (MP) 2030 Prediction Session ANN results

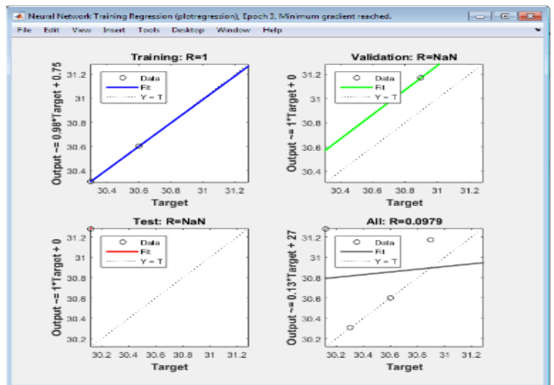

Fig -4: Indore (MP) 2030 Prediction Session ANN results.

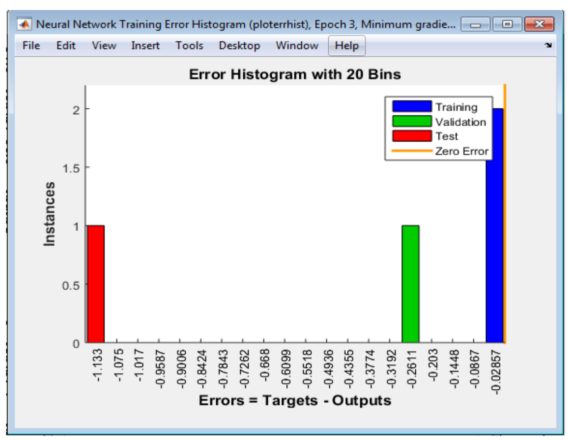

Fig -5: Indore (MP) 2030 Prediction Session ANN MSE.

6. Indore (MP) 2035 Prediction Session ANN results

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

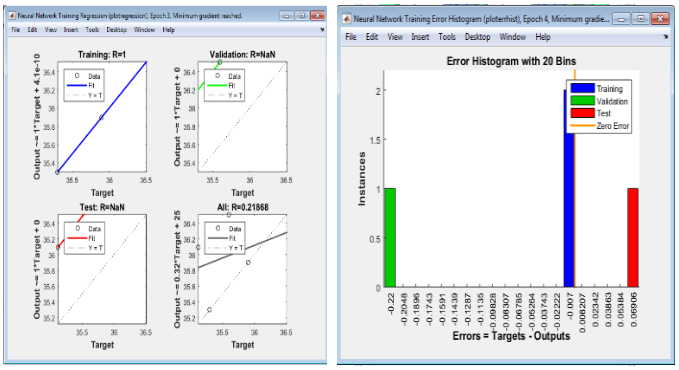

Fig -6: **Indore (MP) 2035 Prediction Session ANN results.**

Table-3: Similarly obtained result in tabulated form Prediction result in MADHYA PRADESH city

Table-4: Similarly obtained result in tabulated form Prediction result in INDIA city

city/year	1999-2000	2004-05	2010-11	2015-16	2025(9) Predict	2030(14) Predict	2035(19) Predict	Accuracy Predict	Accuracy Predict	Accuracy Predict	MSE Predict	MSE Predict	MSE Predict
Mumbai (Mh)	5355	5320	6500	11000	148482	230972	313462	90	90	93	0.015	0.004	-0.007
Delhi i	400	5922	6800	8700	78300	121800	165300	96	49	26	0.007	0.021	0.02
Kolkata (WB)	3692	2653	3670	4000	5949	9254	12559	18	no	no	0.018	0.01	0.06
Chennai (TN)	3124	3036	4500	5000	14706	22876	434644	no	11	24	0.011	0.0067	-0.0015
Ahmedabad (Guj)	1683	1302	2300	2500	15678	24388	33098	no	86	no	-0.012	0.0064	0.0053

the optimal route for collection and disposal of solid waste. It deduces the use integer linear programming model (ILP), GIS-based optimized using ArcGIS Network Analyst tool applied

City/ year	1999-2000	2004-05	2010-11	2015-16	2025 (9) Predict	2030 (14) Predict	2035 (19) Predict	2025 Accuracy Predict	2030 Accuracy Predict	2030 Accuracy Predict	2025 MSE Predict	2030 MSE Predict	2035 MSE Predict
Indore (MP)	350	557	720	850	7569	11774	15979	93	0.9	21	-0.07	-0.26	-0.024
Bhopal (MP)	546	574	350	700	6534	10164	13794	no	94	72	0.027	0.001	0.0021
Jabalpur (MP)	148	216	400	550	4716	7112	9652	49	57	64	-0.0096	-0.011	-0.1294
Gwalior (MP)	10	178	285	300	4743	4578	6213	no	57	29	0.194	0.2699	0.04
Dewas (MP)	258	136	235	436	2286	3682	4997	74	54	no	0.004	0.003	0.011

9. CONCLUSION AND FUTURE SCOPE

Our study on machine learning algorithms for intelligent applications on waste management opens several research issues in the area. Thus, in this section, we summarize conclusions that have potential research opportunities and future directions. The study critically application of machine learning models in the domains of waste management. It is also evident from the analysis that most of the articles used ANN machine learning models. More specifically, the application of ANN was mostly for prediction of waste generation. The evolution and application of artificial intelligence systems in providing innovative solutions to societal challenges. To obtained the results of in terms of Accuracy and improvement of its values up to 96%. They provide recommendations and insights for decision makers for possible actions. This dynamic shift builds on the future application of machine learning models that unlock greater possibilities in solid waste management.

A successful machine learning model heavily depends on data and the performance of the learning algorithms. For optimal results learning algorithms then need to be trained through the collected real-world data and knowledge related to the target application before the system can assist with intelligent decision-making. Based on the review above, machine learning has been applied to solve solid waste challenges in the three domains. The study recommends to use ML models in finding on variables such as cost, route distance and number of trucks, gives the best results. Further research will be carried out in future to realize and validate the tool.

REFERENCES

1. Sorg, T.J. Industrial solid waste management. J. Water Pollut. Control Fed. 1972, 44, 1107–1108.

2. Nai, C.; Tang, M.; Liu, Y.; Xu, Y.; Dong, L.; Liu, J.; Huang, Q. Potentially contamination and health risk to shallow groundwatercaused by closed industrial solid waste landfills: Site reclamation evaluation strategies. J. Clean. Prod. 2021, 286, 125402. [CrossRef]

3. Yang, Z.; Chen, H.; Du, L.; Lu, W.; Qi, K. Exploring the industrial solid wastes management system: Empirical analysis offorecasting and safeguard mechanisms. J. Environ. Manag. 2021, 279, 111627. [CrossRef]

4. Lima, D.Q.; Oliveira, L.C.A.; Bastos, A.R.R.; Carvalho, G.S.; Marques, J.G.S.M.; Carvalho, J.G.; de Souza, G.A. Leather IndustrySolid Waste as Nitrogen Source for Growth of

Common Bean Plants. Appl. Environ. Soil Sci. 2010, 2010, 703842. [CrossRef]

5. Pablos, J.M.; Sichieri, E.P.; Izeli, R.L. Reuse of the industrial solid waste, consisting for foundry sands, in the confection of massivebricks and decorative pieces. J. Archit. Urban. 2009, 0, 112–125.

6. Martinez-Avila, O.; Munoz-Torrero, P.; Sanchez, A.; Font, X.; Barrena, R. Valorization of agro-industrial wastes by producing2-phenylethanol via solid-state fermentation: Influence of substrate selection on the process. Waste Manag. 2021, 121, 403–411.[CrossRef]

7. Krishnan, S.; Zulkapli, N.S.; Kamyab, H.; Taib, S.M.; Din, M.F.B.M.; Majid, Z.A.; Chaiprapat, S.; Kenzo, I.; Ichikawa, Y.; Nasrullah,M.; et al. Current technologies for recovery of metals from industrial wastes: An overview. Environ. Technol. Innov. 2021, 22,101525. [CrossRef]

8. Krishna, R.S.; Mishra, J.; Meher, S.; Das, S.K.; Mustakim, S.M.; Singh, S.K. Industrial solid waste management through sustainablegreen technology: Case study insights from steel and mining industry in Keonjhar, India. Mater. Today Proc. 2020, 33, 5243–5249.[CrossRef]

9. Mavroulidou, M.; Shah, S. Alkali-activated slag concrete with paper industry waste. Waste Manag. Res. 2021, 39, 466–472.[CrossRef]

10. Khan, M.M.H.; Havukainen, J.; Horttanainen, M. Impact of utilizing solid recovered fuel on the global warming potential ofcement production and waste management system: A life cycle assessment approach. Waste Manag. Res. 2021, 39, 561–572.[CrossRef]

11. Viczek, S.A.; Aldrian, A.; Pomberger, R.; Sarc, R. Origins of major and minor ash constituents of solid recovered fuel forco-processing in the cement industry. Waste Manag. 2021, 126, 423–432. [CrossRef] [PubMed]

12. Liu, Y.; Hu, X.; Luo, X.; Zhou, Y.; Wang, D.; Farah, S. Identifying the most significant input parameters for predicting districtheating load using an association rule algorithm. J. Clean. Prod. 2020, 275, 122984. [CrossRef]

13. Ayeleru, O.O.; Fajimi, L.I.; Oboirien, B.O.; Olubambi, P.A. Forecasting municipal solid waste quantity using artificial neuralnetwork and supported vector machine techniques: A case study of Johannesburg, South Africa. J. Clean. Prod. 2021, 289, 125671.[CrossRef]

14. Christensen, T.H.; Damgaard, A.; Levis, J.; Zhao, Y.; Bjorklund, A.; Arena, U.; Barlaz, M.A.; Starostina, V.; Boldrin, A.; Astrup, T.F.;et al. Application of LCA modelling in integrated waste management. Waste Manag. 2020, 118, 313–322. [CrossRef]

15. Guo, H.N.; Wu, S.B.; Tian, Y.J.; Zhang, J.; Liu, H.T. Application of machine learning methods for the prediction of organic solidwaste treatment and recycling processes: A review. Bioresour. Technol. 2021, 319, 124114. [CrossRef]