Solid Waste Management Prediction Using Mining: A Review

Deepika Sharma
Research scholar, Department of Computer Science,SSSUTMS Sehore

Dr.Jitendra Sheetlani

Department of Computer Application,SSSUTMS Sehore

Narendra Sharma
Department of Computer Science & Engg, SSSUTMS Sehore

Abstract - A sustainable Municipal solid waste management is a very big issue for every country. Most of the countries are facing the problem of solid waste management to a long time. Government makes many policies for proper waste management and spends a huge amount of money for waste management. The quantity of generating solid waste increase varies quickly as the rate of the population grows. A sustainable municipal solid waste (MSW) management is a major concern to all governments to protect human health, the environment and to preserve natural resources. For accurate estimation of future waste generation quantities, the design ,planning and operations will be planned very carefully because these are directly affected the human being and environment also. For well planning and strategy, data mining play very important role. The main aim of this paper to present a systematic review of data mining technologies in solid waste management.

Keywords:- Municipal Solid Waste, Data Mining, Waste Management Techniques, Artificial Intelligence, Machine Learning.

1.Introduction :-

Development of a sustainable Municipal Solid Waste Management (MSWM) plan is a difficult process. For efficient MSWM plan and prediction of proper Solid Waste (SW) generation is very much prerequisites. Municipal Solid Waste (MSW) prediction cannot be done directly and depends on so many parameters. Due to suspicions and unavailability of sufficient data, modeling methods are very helpful for prediction of MSW generation. Many researchers have predicted solid waste generation using various modeling methods[1].

The rate of producing solid waste in India is changing rapidly as the population grows. Now, these days working on the level of waste segmentation changes, the conformation of the waste generated evolves, and the technologies available to collect and process waste improve very rapidly [2]. Many new technologies have been discovered for dealing with workable solid waste management. And for the expansion of these technologies, new data mining techniques can be very useful . It can be provided with useful information on decision-makers to make accurate decisions to increase the effectiveness of the decision-making procedure [3].

Municipalities collect about 70-80 % too large Indian cities, whereas small towns and cities collect less than 60% of the waste produced . To open landfills sits and local dump sits more than 90% solid waste has been disposed of. The waste which is remaining after the transportation of waste is openly burned. The quantity of this type of MSW is nearly about 2%, and approximately bout 10% of the collected MSW is burn openly or on fires on landfills. The downside to the open burning of MSW] releases more than 25,000 tons of dangerous gases into big cities in India. India's unceremonious recycling sector gathers most recyclable waste and handovers to Urban Local Bodies to proper disposal [4]. The recognized sectors ceremoniously detach more than 25% of recyclables composed wastes at MSW collection stations and dumps. Even though this number does exclude the measure of reusing before formal collection, it contrasts genuinely well, and the best-reusing rates accomplished around the Informal reusing framework are recently accepting by all the countries of the world, and it is most useful for in SWM in developing countries [5].

In India, non-administrative and government organizations are wanted to set up the segment present in various areas, and to help build up a co-operated small framework it into the By the Ministry of Environment and Forests has been ingrates 'Plastic Waste Management and Handling Rules, 2011'a advance for this purpose. These standards order ULBs to facilitate all MSWM administration parameters, which incorporates waste collectors. The administration makes and attempts legitimate metropolitan strong burn through administration procedures every once in a while. MSW has gone over initial disappointments due to the non-presence of legitimate assets. For overseeing and supervision of MSW, which give a structure to a social event, dealing with and transfer of MSW Indian government has been readied rules and arrangement after 2000 and control them [6].

The government makes and tries proper municipal solid waste management techniques from time to time. MSW has come across initial failures due to the non-existence of appropriate resources. For managing and supervision of MSW, through provides a structure for gathering, handling, and treating MSW. The Indian government has been prepared rules and policy after 2000 and regulates them. New Municipal solid waste rules 2002 include whole. These things pledgee the correctness of present technologies according to Indian circumstances, do not effectively study in past time, expressly concerning the supportable controlling of all types of MSW, and decrease its harmful effects on ecological and human health [7].

2. Data Mining Approach

It is proficient using a training data set to produce a model. When new input data is familiarized to the DM technique, it makes estimate on the beginning of the model. The calculation is appraised for exactness and if the accurateness is adequate, the data mining technique is arranged. If the exactness is not adequate, the data mining technique is competent again and again with an enlarged training data set.[8] This is just a very high-level sample as there are various factors and other steps involved.

2.1 Supervised Learning

Supervised Learning is the furthermost popular paradigm for accomplishment data mining & machine learning operations. It is generally used for data where there is an accurate mapping between input-output data. The dataset, in this situation, is labeled, meaning that the algorithm categorizes the topographies unambiguously and transmits out estimates or cataloguing consequently. [9] As the training period improvements, the algorithm is able to ascertain the interactions between the two variables such that we can predict a new conclusion. Resulting Supervised learning algorithms are task-oriented. As we deliver it with more and examples, it is intelligent to learn more appropriately so that it can commence the task and vintage us the output more truthfully. Some more of the algorithms that originated under managed learning are as follows: Linear regression, random forest, support vector machine, artificial intelligence [9], etc.

There are two foremost types of supervised learning problems: they are arrangement that encompasses expecting a class label and worsening that encompasses expecting a numerical value.

- Classification: Supervised learning problematic that involves envisaging a class label.

- Regression: Supervised learning problem that comprises predicting a numerical label.

Both classification and deterioration difficulties may have one and more input variables and input variables may be any data type, such as numerical or categorical.

2.2 Unsupervised Learning

Unsupervised learning holds the enhancement of practicality bright to work with unlabeled data. This means that human labor is not obligatory to make the dataset machine-readable, consenting much larger datasets to be controlled on by the program. The model learns concluded observation and finds arrangements in the data. Once the model is particular a dataset, it repeatedly finds patterns and relationships in the dataset by generating clusters in it.[10]

In supervised learning, the labels permit the algorithm to find the particular nature of the relationship between any two data points. Nevertheless, unsupervised learning does not have labels to work off of, resulting in the conception of hidden arrangements. Relationships between data points are supposed by the algorithm in an abstract manner, with no input mandatory from human beings. The creation of these unknown structures is what makes unsubstantiated learning algorithms versatile. Instead of a demarcated and set delinquent statement, unsupervised learning algorithms can adapt to the data by dynamically changing hidden buildings.[11] This suggestions more post-deployment expansion than supervised learning algorithms. What it cannot do is add tags to the cluster, like it cannot say this a assemblage of apples or mangoes, but it will dispersed all the apples from mangoes. Suppose we accessible images of apples, bananas and mangoes to the model, so what it does, based on several patterns and interactions it creates clusters and splits the dataset into those clusters. Now if a new data is suckled to the model, it adds it to one of the produced clusters. The sample of unsupervised learning is k-mean clustering, principle section analysis, SVD, FP-growth etc.

There are countless types of unsupervised learning, while there are double main problems that are often encountered by a practitioner: they are clustering that comprises outcome groups in the data and concentration guesstimate that encompasses brief the distribution of data.[20]

- Clustering: Unsupervised learning delinquent that implicates definition groups in data.

- Density Estimation: Unsupervised learning difficult that implicates summarizing the circulation of data.

A. K-Means Clustering

K-means clustering is another basic technique often used in machine learning. While machine learning is often thought of as a fairly new concept, the fundamentals have been around for much longer than many would expect. Specifically, the k-means clustering algorithm has been around since 1967 when it was first developed by a researcher named James MacQueen. Unlike many other machine learning techniques, k-means is used on unlabeled numerical data rather than data that is already defined, making it a type of unsupervised learning. It is one of the most popular unsupervised learning techniques due to its simplicity and efficiency, helping us data scientists out when we don't have the most organized data set. The k-means clustering algorithm assigns data points to categories, or clusters, by finding the mean distance between data points. It then iterates through this technique in order to perform more accurate classifications over time. Since you must first start by classifying your data into k categories, it is essential that you understand your data well enough to do this.[12]

B. Neural Network

Neural networks are one of the learning algorithms used within machine learning. They consist of different layers for analyzing and learning data. Every hidden layer tries to detect patterns on the picture. When a pattern is detected the next hidden layer is activated and so on. The picture of the Audi A7 above illustrates this perfectly. The first layer detects edges. Then the following layers combine other edges found in the data, ultimately a specified layer attempts to detect a wheel pattern or a window pattern. Depending on the amount of layers, it will be or not be able to define what is on the picture, in this case a car. The more layers in a neural network, the more is learned and the more accurate the pattern detection is. Neural Networks learn and attribute weights to the connections between the different neurons each time the network processes data. This means the next time it comes across such a picture, it will have learned that this particular section of the picture is probably associated with for example a tire or a door.[13]

2.3 Reinforcement Learning

Reinforcement learning unswervingly takes motivation from how human actualities learn from data in their lives. It topographies an algorithm that progresses upon itself and learns from new conditions using a trial-and-error method. Satisfactory outputs are stimulated or 'reinforced', and non-favorable outputs are discouraged or 'punished'. Based on the psychological concept of conditioning, buttressing learning works by planting the algorithm in a work environment with an interpreter and an incentive system. In every reiteration of the algorithm, the output result is prearranged to the transcriber, which resolves whether the outcome is favorable or not.[14]

In case of the program definition the truthful solution, the interpreter emphasizes the solution by providing an incentive to the algorithm. If the outcome is not satisfactory, the algorithm is obligatory to restate until it finds a recovering result. In most cases, the incentive system is unswervingly tied to the usefulness of the result.

In typical reinforcement learning use-cases, such as outcome the shortest route between two points on a map, the explanation is not an absolute value. In its place, it takes on a score of usefulness, articulated in a proportion value. The higher this measurement value is, the more reward is prearranged to the algorithm. Thus, the program is trained to elasticity the best thinkable solution for the best imaginable reward. [14] This simple feedback repayment is known as a reinforcement signal.

2.4 Semi-Supervised Learning

In this type of learning, the specified data are an assortment of classified and unclassified data. This amalgamation of labeled and unlabeled data is charity to create an opposite model for the arrangement of data. In most of the conditions, labeled data is unusual and unlabeled data is in wealth (as discussed previously in unsupervised learning description). The target of semi-supervised grouping is to learn a model that will envisage classes of imminent test data better than that from the model engendered by using the considered data alone. The way we learn is comparable to the progression of semi-supervised learning. A child is supplied with:

- Unlabeled data providing by the environment. The environments of a child are full of unlabeled data in the establishment.

- Labeled data from the supervisor. For example, a father explains his children about the names (labels) of objects by pointing toward them and uttering their names[15].

3. Scope of Solid Waste Management

A proper solid waste management technique covers all important parameters like organizational, planning, and economical utilities, elaborate the solution to all complications of the SWM process. These SWM solutions may consist of diverse relationships between all essential factors like development, topography city, and financial side, public health sociology [16].

Generally, peoples think that the SWM system is a straightforward procedure - essentially tapping waste into a vehicle and emptying it at a dump, If this process is simple, then why many cities and metropolises agonize from uncollected waste refuse blocking streets and drains [54]. Sustainable SWM is hardly accomplished without hard effort, new technologies, and learning from past faults. The planning and the executives of a decent solid waste administration framework need contributions from a scope of orders and cautious thought of local conditions [17].

4. CONCLUSION

Clustering and other data mining procedures are very predictable for any analytical work. With the help of analysis, work organization knows about consumer behavior; it can be constructive for increasing the growth of the business. Now that time, data mining is used in every field. In our research work, for refining the quality of solid waste management process,

Unfortunately, although nearly 70% of the waste we produce every day is recyclable, less than a quarter of it is actually recycled. This creates a great need within the waste management industry, a need that can be addressed with policies, outreach and education, and technological improvements.

Many of these tech improvements are beginning to make a noticeable appearance. The use of data mining within the waste management industry has the potential to greatly increase our

capacity to improve our waste management system and recycle more effectively, and could even be the turning point for greater interest in what happens to all of our waste.

The advent of data mining is making the whole process of waste management much easier. In fact, data mining has impacted many industrial and scientific realms in ways that benefit the environment. Data mining is also used to help with route planning garnering better estimates on how much waste is produced where. For instance, waste management organizations use data mining techniques to understand which communities are producing the most waste and targeting them for greater educational outreach on recycling.

References

1. Mladenic, S., Lavraˇc, D., Bohanec, N., Moyle, M.: Data Mining and Decision Support, 1st edn. Springer, USA (2003)
2. Sharma, N., Bajpai, A., Litoriya, R.: Comparison the various clustering algorithms of weka tools. Int. J. Emerg. Technol. Adv. Eng. **2**(5), 73–80 (2012)
3. Sharma, N., Litoriya, R.: Incorporating data mining techniques on software cost estimation: validation and improvement. Int. J. Emerg. Technol. Adv. Eng. **2**(3), 301–309 (2012)
4. Sharma, N.: An analytical study on the importance of data mining for designing. J. Harmon. Res. (JOHR) **7**(2), 44–48 (2019)
5. Heinrichs, J.H., Lim, J.-S.: Integrating web-based data mining tools with business models for knowledge management. Decis. Support Syst. **35**(1), 103–112 (2003)
6. Rok, K., Matjaz, R., Marjan, K.: Integrating data mining and decision support through data mining based decision support system. J. Comput. Inf. Syst. **47**(3), 89–104 (2007)
7. Kantardzic,M.: Data Mining: Concepts, Models,Methods, and Algorithms.Wiley-IEEE Press (2011)
8. Ng, R., Han, J.: Efficient and effective clustering methods for spatial data mining. In: Proceedings of VLDB-94, pp. 144–155 (1994)
9. Litoriya, R., Kothari, A.: Cost estimation of web projects in context with agile paradigm: improvements and validation. Int. J. Softw. Eng. **6**(2), 91–114 (2013)
10. Singh, N.S., Gill, S.: Analysis and study of K-means clustering algorithm. Int. J. Eng. Res. Technol. **2**(7), 2546–2551 (2013)
11. Oyelade, O., Oyelade, J.: Application of k-means clustering algorithm for prediction of students' academic performance. Int. J. Comput. Sci. Inf. Secur. **7**(1), 292–295 (2010)
12. Han, J., Kamber, M., (2006) Data mining, concepts and techniques, Second Edition, Morgan Kaufmann Pub.
13. Kaygulu, M.S. (1999). Supervised and unsupervised learning techniques in data mining, Unpublished Master Thesis, Izmir: Graduate School of Natural and Applied Sciences of DokuzEylul University.
14. Kantardzic, M., (2003) Data Mining: Concepts, Models, Methods, and Algorithms, John Wiley & Sons Inc.
15. R. Ng and J. Han, "Efficient and effective clusteringmethods for spatial data mining", Proc of VLDB-94, 1994.
16. Shakiba Khademolqorania *, Ali Zeinal Hamadania "An Adjusted Decision Support System through Data Mining and Multiple Criteria Decision Making" The 2nd International Conference on Integrated Information 2013.
17. Livani, E., & Raymond Nguyen Jörg (2013). A Hybrid Machine Learning Method and its Application in Municipal Waste Prediction. ICDM 2013: Advances in Data Mining. Applications and Theoretical Aspects, pp. 166-