

## Proposal for a Logistics Regression Model for the Classification of Urban Solid Wastes with Image Processing

V.M. Romero-Juárez\*, M. Castro-Bello\*, V. López-Caballero\*\*, C. Morales-Morales\*, C.V. Marmolejo-Vega\*, S.R. Zagal-Barrera\*, A. Bárcenas-Nava\*

\*Tecnológico Nacional de México/Instituto Tecnológico de Chilpancingo, Av. José Francisco Ruiz Massieu, No. 5, Fracc. Villa Moderna, CP. 39090, Chilpancingo, Guerrero.

\*\*Tecnológico Nacional de México/CENIDET, Interior Internado Palmira S/N, Col. Palmira, C.P. 62490, Cuernavaca, Morelos.

**Abstract:** The environmental problem of Urban Solid Waste Management (USWM) has been increasing exponentially in recent years, this situation is expected to worsen. This article proposes the development of a Robust Logistics Regression Model (RLRM) for the classification of Urban Solid Waste (USW) in organic and inorganic, using convolutionary neural networks that minimize atypical data sensitivity, as well as analysis of parameters and validation of the model that contributes to decision-making.

**Keywords:** urban solid waste, convolutionary neural networks, logistical regression, model, parameters.

### 1. INTRODUCCIÓN

En el marco de la sustentabilidad ambiental y de los procesos de urbanización, privatización y descentralización; la gestión de los Residuos Sólidos Urbanos (RSU) actualmente constituye una preocupación de singular importancia por sus impactos directos e indirectos, algunos de ellos irreversibles y permanentes, tanto en el medio ambiente (aire, agua, tierra, paisaje) como la salud. El efecto ambiental más evidente del manejo inadecuado de los RSU, son el deterioro estético de las ciudades, del paisaje natural, urbano, rural, con la consecuente devaluación, tanto de los predios donde se localizan los vertederos como de las áreas vecinas por el abandono y la acumulación de la basura, sin embargo, entre los efectos ambientales más severos están la contaminación del suelo, del aire y los cuerpos de agua, ocasionada por el vertimiento directo de los residuos (Rondon-Toro *et al.*, 2016).

En los últimos años la tecnología y la modelación matemática han avanzado de manera considerable y sus aplicaciones en áreas del conocimiento, como en las ciencias ambientales, salud, educación, seguridad, entre otras.

Investigaciones en Golbaz *et al.* (2018), han desarrollado modelos de Regresión Lineal Múltiple (MLR) con aprendizaje automático basados en neuronas y algoritmos de Kernel para predecir RSU en hospitales; así como modelos de regresión multivariable para pronosticar la cantidad de RS (Al-Salem *et al.*, 2018). Por otra parte, Ayeleru *et al.* (2018) determinaron la caracterización y el patrón de los RSU, mediante el software Minitab para realizar la previsión estadística basada en datos históricos e indicadores de pronósticos, además se usó el modelo de tendencia lineal para el análisis de tendencia temporal; y Kumar *et al.* (2021) proponen una solución automatizada basada en Inteligencia Artificial (IA) para clasificar las corrientes de residuos médicos relacionados con el COVID mediante histogramas en escala de grises,

propiedades del histograma estadístico de primer orden, función de Patrón Binario Local (PBL), esquema de fusión a nivel de decisión.

La población mundial genera aproximadamente 2,010 millones de toneladas de RSU anualmente, al menos el 33% de esa cantidad no se gestiona de manera ambiental y segura. Para 2030, se espera que se generen 2,590 millones de toneladas al año, y para el 2050, se alcancen los 3,400 millones (Kaza *et al.*, 2018). En América Latina la clasificación de los RSU se realiza de manera manual con el riesgo de que operarios contraigan enfermedades como tífus, cólera, sarnas, hepatitis, entre otras, lo que sugiere la necesidad de desarrollar tecnologías con la aplicación de modelos matemáticos para optimizar tiempos y eficiencia en la clasificación de los RSU (Sáez y Urdaneta, 2014).

A falta de estrategias para mejorar la protección y cuidado del medio ambiente, diversas investigaciones han aplicado metodologías para clasificar RSU con procesamiento de imágenes y RNC, no obstante, de tener buenos resultados, son computacionalmente muy costosas, carecen de interpretación de covariables, generan sobreajuste, son sensibles a datos heterogéneos y les falta estimación y análisis de parámetros (Clavijo *et al.*, 2022).

Esta investigación propone el diseño y desarrollo de un modelo estadístico probabilístico de regresión logística con procesamiento de imágenes de RSU aplicando redes neuronales convolucionales con la finalidad de contribuir al cuidado del medio ambiente.



### 1.1. Trabajos relacionados

Trabajos relacionados como el de Huang *et al.* (2020) proponen un modelo de clasificación combinado basado en tres modelos CNN preentrenados (VGG19, DenseNet169 y NASNetLarge) para procesar la base de datos ImageNet. Mao *et al.* (2021) aplicaron el modelo de red neuronal convolucional (CNN) DenseNet121 para el reconocimiento de imágenes, y para la evaluación del rendimiento utilizaron el conjunto de datos TrashNet compuesto por un total de 2,527 imágenes con seis categorías de residuos diferentes, además del aumento de datos, y se plantearon como objetivo utilizar el algoritmo genético (GA).

Otro proyecto, es el de Hurst *et al.* (2022) donde utilizaron un conjunto de datos de clasificación de residuos para evaluar el rendimiento de una red neuronal convolucional de cinco capas, la cual se entrenó con dos resoluciones de imagen diferentes.

En Brasil se desarrollaron modelos de predicción estadística para el valor calorífico basados en datos experimentales tanto con el uso de la composición gravimétrica y/o el análisis elemental/próximo y el contenido de la humedad de los residuos recogidos (Drudi *et al.*, 2019). Clavijo *et al.* (2022) en la universidad EAFIT de Medellín, Colombia realizaron dos modelos de clasificación con base en la regresión logística, en su versión clásica y en su versión robusta, donde se aplicó el coeficiente de correlación de Kendall, el coeficiente de correlación de Pearson y aplicaron el algoritmo de detección de bordes de Canny para el suavizado de imágenes, obtención de gradientes, supresión de falsos máximos y el umbral de histéresis.

Y en México Salazar y Filemón (2020), desarrollaron una aplicación para la identificación de botellas PET, papel y cartón, en el que se implementó el uso de una Red Neuronal Convolutiva (RNC).

## 2. METODOLOGÍA

Para efectos de este trabajo, se usará el conjunto de imágenes del repositorio trashnet, el cual abarca seis categorías: vidrio, papel, carton, metal y basura.

Actualmente, el conjunto de datos consta de 2,527 imágenes: 501 vidrio, 594 papel, 403 cartón, 482 de plástico, 410 metales, y 137 basura, todo el dataset se ha redimensionado a 512 x 384.

Con la finalidad de reducir la sensibilidad de datos atípicos, se aplicarán redes neuronales convolucionales y la ecuación de distribución de probabilidad, ecuación 1, figura 1.

$$y_i = x_i' \beta + \varepsilon_i \quad (1)$$

donde  $x_i' = [1, x_{i1}, x_{i2}, \dots, x_{ik}]$ ,  $\beta' = [\beta_0, \beta_1, \beta_2, \dots, \beta_k]$ , y la variable de respuesta  $y_i$  toma los valores de 0 (orgánico) o 1 (inorgánico). Se supondrá que la variable  $y_i$  es una variable aleatoria de Bernoulli, cuya distribución de probabilidad es, tabla 1:

**Tabla 1. Distribución de probabilidad.**

$y_i$	Probabilidad
1	$P(y_i = 1) = \pi_i$
0	$P(y_i = 0) = 1 - \pi_i$

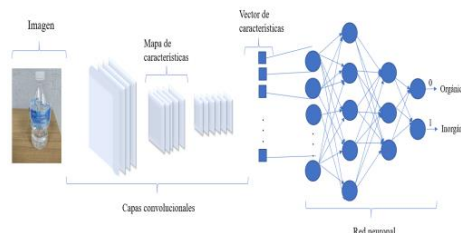


Figura 1. Clasificación de RSU con una red neuronal convolucional.

## 3. RESULTADOS

Los resultados esperados es un MRLR que clasifique RSU en orgánicos e inorgánicos, que minimize el error de estimación y la sensibilidad a datos atípicos, así como la propuesta de una base de datos de RSU orgánicos.

## 4. CONCLUSIONES

El presente trabajo pondrá a disposición un modelo de regresión logística robusto así como la aplicación de una red neuronal convolucional para la clasificación de RSU en orgánicos e inorgánicos.

## REFERENCIAS

- Alejandra Vélez Clavijo, M., Moreno Ramírez, V., Palacio Jaramillo, A., José Wilches Rivas Asesor, J., Laniado Rodas, H., 2021. Regresión logística robusta y generación de datos sintéticos para la clasificación de residuos sólidos. <https://doi.org/10.13140/RG.2.2.18263.55204>
- Al-Salem, S.M., Al-Nasser, A., Al-Dhafeeri, A.T., 2018. Multi-variable regression analysis for the solid waste generation in the State of Kuwait. *Process Safety and Environmental Protection* 119. <https://doi.org/10.1016/j.psep.2018.07.017>
- Ayleru, O.O., Okonta, F.N., Ntuli, F., 2018. Municipal solid waste generation and characterization in the City of Johannesburg: A pathway for the implementation of zero waste. *Waste Management* 79. <https://doi.org/10.1016/j.wasman.2018.07.026>
- Drudi, K.C.R., Drudi, R., Martins, G., Antonio, G.C., Leite, J.T.C., 2019. Statistical model for heating value of municipal solid waste in Brazil based on gravimetric composition. *Waste Management* 87, 782–790. <https://doi.org/10.1016/J.WASMAN.2019.03.012>
- Golbaz, S., Nabizadeh, R., Sajadi, H.S., 2019. Comparative study of predicting hospital solid waste generation using multiple linear regression and artificial intelligence. *Journal of Environmental Health*

- Science and Engineering 17.  
<https://doi.org/10.1007/s40201-018-00324-z>
- Huang, G.L., He, J., Xu, Z., Huang, G., 2020. A combination model based on transfer learning for waste classification. *Concurr Comput* 32, e5751.  
<https://doi.org/10.1002/CPE.5751>
- Hurst, W., Ebo Bennin, K., Kotze, B., Mangara, T., Nnamoko, N., Barrowclough, J., Procter, J., 2022. Solid Waste Image Classification Using Deep Convolutional Neural Network. *Infrastructures* 2022, Vol. 7, Page 47, 47.  
<https://doi.org/10.3390/INFRASTRUCTURES7040047>
- Kaza, S., Yao, L.C., Bhada-Tata, P., Van Woerden, F., 2018. What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050.  
<https://doi.org/10.1596/978-1-4648-1329-0>
- Kumar, N.M., Mohammed, M.A., Abdulkareem, K.H., Damasevicius, R., Mostafa, S.A., Maashi, M.S., Chopra, S.S., 2021. Artificial intelligence-based solution for sorting COVID related medical waste streams and supporting data-driven decisions for smart circular economy practice. *Process Safety and Environmental Protection* 152.  
<https://doi.org/10.1016/j.psep.2021.06.026>
- Mao, W.L., Chen, W.C., Wang, C.T., Lin, Y.H., 2021. Recycling waste classification using optimized convolutional neural network. *Resour Conserv Recycl* 164, 105132.  
<https://doi.org/10.1016/J.RESCONREC.2020.105132>
- Montgomery, D.C., Peck, E.A., Vining, G.G., 2006. *Introducción al Análisis de Regresión Lineal*, Syria Studies.
- Rondon Toro, E., Szanto Narea, M., Pacheco, J.F., Contreras, E., Galvez, A., 2016. *Guía General para la gestión de residuos sólidos domiciliarios*. Manuales de la CEPAL.
- Sáez, A., Urdaneta, J.A., 2014. *Manejo de residuos sólidos en América Latina y el Caribe*, Omnia Año.
- Repositorio Institucional del Tecnológico Nacional de México (RI - TecNM): Desarrollo de una aplicación para la identificación de PET, cartón y papel por medio de una red convolutiva [en línea], (sin fecha). Repositorio Institucional del Tecnológico Nacional de México (RI - TecNM): Home. [Consultado el 11 de noviembre de 2023]. Disponible en:  
<https://rinacional.tecnm.mx/jspui/handle/TecNM/2688>
- GitHub - garythung/trashnet: Dataset of images of trash; Torch-based CNN for garbage image classification [en línea], (sin fecha). GitHub. [Consultado el 11 de noviembre de 2023]. Disponible en:  
<https://github.com/garythung/trashnet>