

Clasificación y Análisis de emisiones de Gases de Efecto Invernadero

Helena Patricia Carrillo Soto*

Abstract

The abstract is a brief (usually one paragraph) summary of the whole paper, including the problem, the method for solving it (when not obvious), the results, and the conclusions suggested or drawn. Do not write the abstract as a hasty afterthought. Profesor si está leyendo esto aun no escribo un abstract porque aun no tengo terminado el análisis.

Palabras clave: Gases de Efecto Invernadero, Aprendizaje No Supervisado, Machine Learning

1 Introducción

Los Gases de Efecto Invernadero (GEI) son los componentes de la atmósfera que absorben la radiación emitida por la superficie de la Tierra y la emiten de regreso.

Los principales GEI son el vapor de agua (H_2O), el dióxido de carbono (CO_2), el óxido nitroso (N_2O), el metano (CH_4) y el ozono (O_3).

También se encuentran GEI creados en su totalidad por el ser humano, como los halocarbonos (CFCs, HCFCs, HFCs y PFCs) los cuales contienen cloro, bromo o flúor y carbono. Estos compuestos son una de las causas del agotamiento de la capa de ozono en la atmósfera.

Los Gases de Efecto Invernadero están directamente relacionados con el calentamiento global y este a su vez con el cambio climático.

Este tema ha cobrado relevancia en los últimos años con el rápido aumento de la temperatura media global, los intentos fallidos de distintos gobiernos de reducir sus emisiones y la llamada de científicos alrededor del mundo de tomar medidas urgentes para controlar el cambio climático.

*Facultad de Ciencias Fisico Matemáticas, Universidad Autónoma de Nuevo León (UANL). Email: helenacarrilloso@uanl.edu.mx

2 Metodología y Datos

En el presente trabajo se utilizarán métodos de Machine Learning para analizar los datos correspondientes al Inventario Nacional de Emisiones de Gases y Compuestos de Efecto Invernadero del INECC que se puede encontrar en la siguiente liga .

Los datos consisten en las emisiones anuales totales, y desglosadas por industria, de los diferentes gases de efecto invernadero para los años de 1990 a 2021.

3 Aprendizaje No Supervisado

El aprendizaje no supervisado, usa algoritmos de machine learning para analizar y clasificar datos no etiquetados.

Se utilizan principalmente para encontrar patrones ocultos en los datos.

3.1 K-Means

El algoritmo K-means es un método popular de clustering por aprendizaje no supervisado.

El algoritmo se basa en agrupar los datos en k grupos en base a su media.

En él asumimos que se tiene un set de objetos $D = \{ p_1, p_2, \dots, p_n \}$ donde cada objeto $p_i = \{ p_{i1}, p_{i2}, \dots, p_{in} \}$ representa un punto en el espacio R^m donde m es el número de atributos o dimensiones de los objetos del set D .

Se define k como el número de clusters y m_k la media del cluster C_k . El objetivo general del método es minimizar la función criterio $e(k)$.

Normalmente la función criterio que se minimiza es la suma de cuadrados del error:

$$e(k) = \sum_{i=1}^k \sum_{p_j \in C_i} dist(p_j, m_i)^2$$

donde

p_j - punto en el espacio R^m que reopresenta el punto p_j .

m_i - media del cluster C_i .

$dist(p_j, m_i)$ - distancia Euclidiana del punto p_j a la media (centro) del cluster más cercano C_i .

3.2 K-means en el análisis de GEI

El algoritmo de K-means se ha usado para el análisis de emisiones de Gases de Efecto invernadero en múltiples ocasiones.

Anna Kijewska y Anna Bluszcz hicieron uso de él en 2016 para analizar los niveles de emisiones de GEI variantes en países de Europa. (Kijewska & Bluszcz, 2016)

En 2018, Lozza y Bellini lo utilizaron para clasificar sistemas ganaderos para estimaciones de GEI. (Lozza & Bellini Saibene, 2018)

En esta ocasión se usará el algoritmo para categorizar las actividades que aportan al total de emisiones de Gases de Efecto Invernadero en México de los años de 1990 a 2021.

3.3 Resultados análisis K-means

Para establecer el número de clusters a utilizar se utilizó el método del codo en el cuál se calcula la suma de las distancias al cuadrado para diferentes valores de k y se busca el valor de k donde la disminución en la suma se realentiza o se acerca a su límite.

El resultado obtenido fue el siguiente:

Para tener un criterio más claro se puede calcular la distancia entre la recta creada por el mínimo y máximo número de clusters y sabemos que se optimiza k en el número que tenga una mayor distancia a la recta:

En base a esto se usan 4 clústers.

Los resultados arrojan los siguientes centros:

TABLE 1: Centros: Puntos donde se encuentran los centros de cada uno de los clusters C_k .

C_k	x	y	z	w	v	u	r	s
1	-0.0907	-0.0822	-0.0089	0.0020	-0.0469	0.0019	0.0019	-0.1110
2	-0.1175	10.6739	-0.1918	-0.0760	-0.0551	-0.0724	-0.0724	3.0160
3	5.7762	-0.1407	0.6824	-0.0760	-0.0551	-0.0724	-0.0724	5.4903
4	-0.1129	-0.1810	-0.1918	-0.0760	19.3721	-0.0724	-0.0724	-0.1575

También se obtienen los tamaños de grupo:

TABLE 2: Tamaño: Tamaño de cada uno de los clusters C_k .

Cluster	nk
1	4022
2	32
3	64
4	10

Debido a que el espacio es de 8 dimensiones solo se muestran algunas caras:

4 Apendizaje Supervisado

A diferencia del aprendizaje no supervisado, el aprendizaje supervisado utiliza datos etiquetados. Este utiliza datos de entrenamiento para enseñar a los modelos a generar la salida deseada.

4.1 Modelos de aprendizaje supervisado en el análisis de GEI

Distintos modelos de aprendizaje supervisado se han usado a lo largo del tiempo para analizar emisiones de Gases de Efecto Invernadero.

4.2 Análisis y Resultados

You can make lists with automatic numbering . . .

1. Like this,
2. and like this.

. . . or bullet points . . .

- Like this,
- and like this.

References should be in APA style. Examples are below.

References

- Kijewska, A. & Bluszcz, A. (2016). Research of varying levels of greenhouse gas emissions in European countries using the k-means method. *Atmospheric Pollution Research*, 7(5), 935–944, <https://doi.org/https://doi.org/10.1016/j.apr.2016.05.010> <https://www.sciencedirect.com/science/article/pii/S1309104215305754>.
- Lozza, A. & Bellini Saibene, Y. (2018). Clasificación de sistemas ganaderos para estimación de gases de efecto invernadero. In *Conferencia Latinoamericana sobre Uso de R en Investigación+ Desarrollo (LatinR 2018)-JAIIO 47 (CABA, 2018)*.