

Estimación de Densidad una tarea de Aprendizaje Automático Density Estimation: A Machine Learning Task

Daniel Cantón Enríquez ¹, Jorge-Luis Pérez-Ramos ¹, Selene Ramírez-Rosales ¹, Luis-Antonio Díaz-Jiménez ¹ and Carlo-Giovanni Cetina-Camacho ¹

¹Centro de Investigación e Innovación en Ciencias de la Computación y Tecnología Educativa, Facultad de Informática,
Universidad Autónoma de Querétaro,
Av. de las Ciencias S/N, Juriquilla, Santiago de Querétaro, Qro., 76230. México
daniel.canton@uaq.mx

Fecha de recepción: 2 de octubre de 2025

Fecha de aceptación: 13 de marzo de 2026

Resumen. La estimación de densidad busca aproximar una función de densidad de probabilidad usando un conjunto de datos. Asimismo, la estimación de parámetros es el proceso de calcular los valores de los parámetros de un modelo de densidad. En este trabajo, se tiene como objetivos el estudio de las técnicas de estimación de densidad. Así como su aplicación en dos casos de uso: (i) un análisis de frecuencias para nuevos casos confirmados de COVID-19 usando una distribución Gumbel; y (ii) el desarrollo de un modelo de sustracción de fondo adaptable para la detección de movimiento en un ambiente interior. Los resultados obtenidos demuestran que las técnicas de estimación de densidad pueden ser aplicados en la solución de diferentes problemas. Por último, se proponen diversos trabajos a futuro centrados en tres aspectos principales.

Palabras clave: Análisis de Frecuencias, Inteligencia Artificial, Modelos de Univariados, Modelos Multivariados, Reconocimiento de Patrones.

Summary. Density estimation seeks to approximate a probability density function using a dataset. Parameter estimation is the process of calculating the parameter values of a density model. This work aims to study density estimation techniques and their application in two use cases: (i) a frequency analysis for new confirmed COVID-19 cases using a Gumbel distribution; and (ii) the development of an adaptive background subtraction model for motion detection in an indoor environment. The results obtained demonstrate that density estimation techniques can be applied to solve different problems. Finally, several future projects are proposed, focusing on three main aspects.

Keywords: Artificial Intelligence, Frequency Analysis, Multivariate Models, Pattern Recognition, Univariate Models.

1 Introducción

La Inteligencia Artificial (IA) es un área del conocimiento que comienza a mediados del siglo XX, específicamente entre 1940 a 1950. Los aportes iniciales comenzaron con (1) la cibernética de Rober Wiener, el cual describió el control y la estabilidad en redes eléctricas; (2) la teoría de la información de Claude Shannon para describir señales digitales; y (3) la teoría de la computación de Alan Turing que demostró que cualquier forma de computación podía describirse digitalmente.

Sin embargo, la inversión y el interés en la IA se dispararon en las primeras décadas del siglo XXI, cuando el Aprendizaje Automático (AA) se aplicó a muchos problemas en el sector académico y la industria debido al desarrollo de nuevos modelos matemáticos, la aplicación de hardware potente y a la recopilación de inmensos conjuntos de datos. AA trata de diseñar algoritmos que extraen automáticamente información valiosa de los datos [1]. El énfasis en “automático”, es que, el AA se preocupa por las metodologías de propósito general que se pueden aplicar a muchos conjuntos de datos, mientras se genera algo que sea significativo.

Existen tres conceptos que están en el centro del AA: (i) datos, (ii) un modelo y (iii) aprendizaje. Dado que el AA está impulsado por datos, (i) los datos son el núcleo del AA. El objetivo del AA es diseñar metodologías de propósito general para detectar patrones útiles de los datos, de forma ideal sin mucha experiencia en un dominio específico [2]. Por ejemplo, dado un gran corpus de documentos, los métodos de AA se pueden usar para encontrar de forma automática temas relevantes, tales como, resumen, ontología, relaciones, entre otros [3]. Para lograr este objetivo, se diseñan (ii) modelos que normalmente están relacionados con el proceso que genera datos, de forma similar al modelo del conjunto de datos que se proporciona. De acuerdo con [4], el objetivo es encontrar buenos modelos que generalicen bien datos aún no vistos, que puedan ser útiles en el futuro. Por último, el (iii) el

aprendizaje se puede entender como una forma de aprendizaje para encontrar de forma automática patrones y/o estructuras en los datos mediante la optimización de los parámetros del modelo.

Desde otra perspectiva, el AA es un subconjunto de algoritmos dentro del campo de investigación de la IA. Sin embargo, existen un amplio número de métodos y técnicas utilizadas en AA. Tradicionalmente, el AA se puede dividir en tres categorías, que corresponden con el paradigma de aprendizaje. El primero tipo es el aprendizaje supervisado o predictivo, donde el objetivo es aprender una correspondencia de las entradas x a las salidas y , dado un conjunto etiquetado de pares entradas-salidas $D = \{x_i, y_i\}_{i=1}^N$; donde D es el conjunto de entrenamiento y N es el número de ejemplos de entrenamiento. El segundo tipo de aprendizaje es el no supervisado o descriptivo. Aquí solo se tiene las entradas $D = \{x_i\}_{i=1}^N$ y el objetivo es encontrar las estructuras o patrones interesantes en los datos. Esto también es conocido como descubrimiento de conocimiento. Este es un problema mucho menos definido, ya que no se conoce que patrón buscar y no hay métrica de comparación como en el aprendizaje supervisado. Por último, hay un tercer tipo de aprendizaje, conocido como aprendizaje por refuerzo, en el que un agente interactúa secuencialmente en un entorno inicialmente desconocido, pero en el que aprende cómo actuar o comportarse cuando se le dan señales ocasionales de recompensa o castigo.

Por otro lado, los algoritmos de AA se pueden dividir de acuerdo con el problema que se busca resolver. En la Figura 1, se muestran las principales tareas dentro del AA, las cuales pueden variar de acuerdo con los distintos autores. Por ejemplo, [2] solo reconocen cuatro de las seis tareas descritas en la Figura 1, excepto agrupamiento y causalidad. En contraste [1], reconoce todas las tareas descritas, excepto causalidad. Si bien, el problema de asumir o determinar una relación causa-efecto, no es un problema reciente [5].

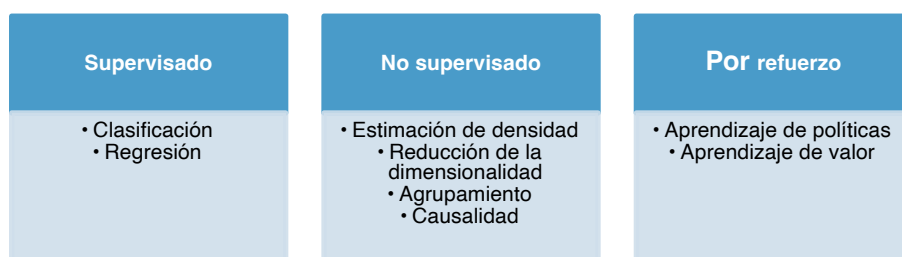


Figura 1. Tareas dentro del aprendizaje automático de acuerdo con el tipo de aprendizaje.

Como se ha revisado, el AA es un campo de investigación extenso, tanto como, por el enfoque de aprendizaje, la problemática a resolver (tareas) y la amplia variedad de aplicaciones en que se pueden trabajar. Esta investigación tiene como objetivos (i) estudiar la tarea de estimación de densidad dentro del AA. Esta investigación tiene como objetivos i) estudiar la tarea de estimación de densidad dentro del AA. Así mismo, se ii) describen dos aplicaciones de estimación de densidad: (a) un análisis de frecuencias aplicado a casos de Covid-19 y (b) un modelo de fondo para la detección de movimiento.

2 Estado del arte

Estimación de densidad: La estimación de densidad es el problema de aproximar la Función de Densidad de Probabilidad (FDP) utilizando un conjunto de datos X . Concretamente, si se observa $x_1, x_2, \dots, x_n$, entonces se recupera la FDP subyacente que genera X . En la Figura 2, se muestra la topología de las técnicas de estimación de densidad. Así mismo, en el ajuste de una FDP existen dos enfoques principales: (i) paramétrico y (ii) no paramétrico. La ventaja de los (a-i) métodos paramétricos unimodales es su simplicidad teórica, que permite modelar los datos con una baja complejidad computacional. No obstante, [6] mencionan que la forma local de una FDP puede ser atípica, compleja o multimodal debido a la naturaleza estocástica de la mayoría de los fenómenos estudiados. Sin embargo, una mezcla de distribuciones paramétricas que considere múltiples modas puede modelar estos retos durante un horizonte de tiempo que sea estacionario. Las mezclas de distribuciones paramétricas más utilizadas son las basadas en Weibull [7-8] y Gaussiana [9-10]. Por otro lado, los (a-ii) métodos de distribución no paramétricos, no asumen una estructura previa en los datos. Un método clásico de este tipo es el histograma, el cual tiene como desventaja definir el número de particiones. Por otro lado, en la literatura el método *Kernel Density Estimation* (KDE) ha sido utilizado principalmente con las funciones núcleo: Gaussiana [5-6, 11-12] y Epanechnikov [13]. No obstante, para ambos enfoques se necesita la estimación de los parámetros que ajustan a una FDP. Si bien, los métodos no paramétricos, como KDE, no necesitan estimar los parámetros de su función núcleo. Si es necesario optimizar el ancho de banda que permite ajustar una FDP a la distribución teórica de los datos.

Estimación de parámetros: La estimación de parámetros que ajustan a una FDP representa un nuevo reto. En la revisión de la literatura, los principales métodos en la estimación de parámetros han sido: *Maximum Likelihood Estimation* (MLE) [7, 12, 14] y *Expectation-Maximization* (EM) en sus diferentes variantes [15-17]. La tarea principal en la estimación de parámetros es maximizar el valor esperado de los datos que pertenecen a una variable compleja. Por lo tanto, la estimación de parámetros es esencialmente un problema de optimización. En consecuencia, se han propuesto técnicas de búsqueda estocásticas para encontrar los parámetros que ajustan a una FDP [8].



Figura 2. Topología de la tarea de estimación de densidad.

3 Materiales y métodos

3.1 Análisis de frecuencias

Un análisis de frecuencias teórico significa ajustar una serie de datos a una Función de Distribución Acumulada (FDA), que representa la probabilidad de ocurrencia de una variable aleatoria. Este procedimiento deberá aplicarse cuando se desee conocer un evento asociado a un período de retorno mayor que la longitud máxima de los datos registrados; por eso se llama teórica porque no es posible estimar el evento usando una tabla de frecuencia empírica. Existen varias FDA, las más utilizadas son las monovariadas monomodales, en especial las distribuciones de valores extremos. También, se utilizan las FDA multivariadas, es decir, pueden tener en cuenta dos o tres conjuntos de datos. Para los datos diarios de Covid-19, se propone utilizar la distribución Gumbel [18]. También, conocida como distribución de valores extremos tipo I. Donde su FDA se describe en la Ecuación 1.

$$F(x) = e^{-e^{-\alpha(x-\beta)}} \quad (1)$$

Donde la Ecuación 2 es

$$\alpha = \frac{1.2825}{s}; \quad \beta = \bar{x} - 0.45s \quad (2)$$

Donde s es la desviación estándar, $\bar{x}$ es la media, α es el parámetro de escala y β es el parámetro de forma. Luego, para igualar la FDP del periodo de retorno $P = 1 - \frac{1}{T}$ con la FDA se tiene la Ecuación 3.

$$F(X) = \int_0^X f(X)dX = P(x \leq X) = 1 - \frac{1}{T} \quad (3)$$

Resolviendo x se tiene la Ecuación 4.

$$x = \beta - \frac{1}{\alpha} \left[\ln \left(\ln \left(\frac{T}{T-1} \right) \right) \right] \quad (4)$$

3.2 Modelo de fondo

En este caso de uso se utiliza un modelo de sustracción de fondo basado en la aproximación a una gaussiana. Esta aproximación modela a cada píxel como una distribución gaussiana, si la intensidad del píxel no se puede modelar por la FDP, se asume que corresponde a un objeto detectado. Los parámetros de cada componente se inicializan: μ con el valor de la primera imagen y σ con un valor preestablecido. Debido al cambio continuo de la variación en la iluminación el modelo estimado es válido por un corto periodo de tiempo. Para lograr un modelo que se adapte a las variaciones de luz, los parámetros de la (μ_1, σ_1) para cada píxel son actualizados. De acuerdo con Jiménez y otros [19], los parámetros se pueden estimar mediante una versión recursiva de EM, que se define con las Ecuaciones (5) y (6).

$$\mu_{k+1} = \rho\mu_k + (1 - \rho)x_k \quad (5)$$

$$\sigma_{k+1}^2 = \rho\sigma_k^2 + (1 - \rho)(x_k - \mu_{k+1})^2 \quad (6)$$

donde $\rho \in [0,1]$, representa el factor de convergencia o la tasa de actualización del modelo. Con los parámetros estimados, se puede definir un criterio de pertenencia, que permita determinar si un píxel corresponde al fondo o a un movimiento. Este criterio está definido por la distancia normalizada, véase Ecuación (7).

$$B = \begin{cases} 1, & \text{si } (x_t - \mu)^2 > \beta\sigma^2 \\ 0, & \text{en otro caso} \end{cases} \quad (7)$$

4 Resultados experimentales

En esta sección, se muestran los resultados de los dos casos de uso desarrollados. Tanto en el análisis de frecuencias como en el modelo de fondo se busca aproximar un FDP de tipo monovariado monomodal. Los datos utilizados para los casos diarios de COVID-19 fueron descargados del repositorio de la Universidad Johns Hopkins. Por otro lado, la secuencia de imágenes utilizada fue tomada de [20], que contiene 1,700 imágenes de una resolución 320x240 píxeles. A continuación, se detallan los resultados obtenidos al aplicar cada modelo de densidad.

4.1 Análisis de frecuencias

A nivel mundial, la enfermedad de coronavirus COVID-19 se propagó rápidamente. En ese contexto de aumentos de contagios, es fundamental conocer la probabilidad de que se notifique un número específico de casos. A continuación, se presentan la probabilidad de incidencia de nuevos casos confirmados en dos países afectados en Latinoamérica: Argentina, México y Brasil. La Tabla 1, muestra la estimación de los parámetros de las distribuciones utilizadas. Este análisis probabilístico proviene del análisis de frecuencia de cada uno de los países, véase las Figuras 3-5.

Tabla 1. Parámetros para el ajuste del modelo Gumbel.

País	N° de datos analizados	Parámetros Gumbel	
		Escala	Forma
Argentina	540	7,448	4,556
México	544	2,498	4,529
Brasil	546	300,052	13,869

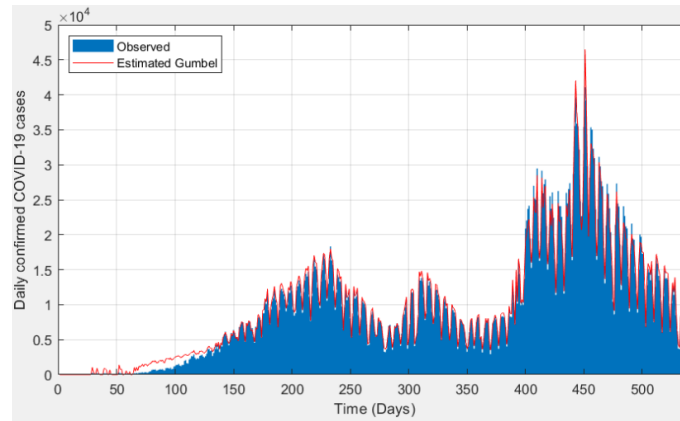


Figura 3. Caracterización Gumbel de nuevos casos confirmados diarios en Argentina.

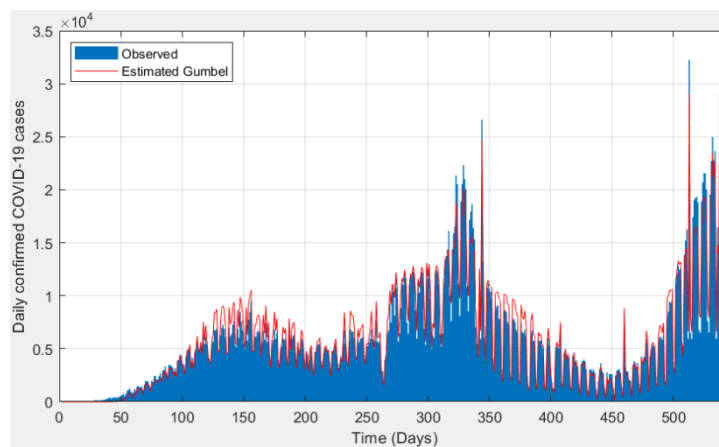


Figura 4. Caracterización Gumbel de nuevos casos confirmados diarios en México.

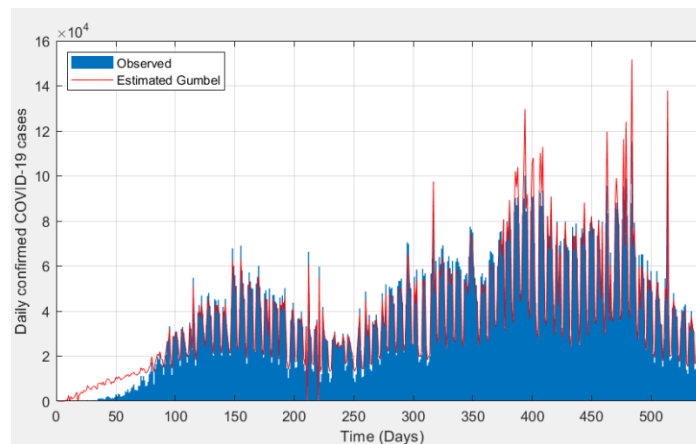


Figura 5. Caracterización Gumbel de nuevos casos confirmados diarios en Brasil.

4.2 Modelo de fondo

Los resultados muestran que la substracción del fondo y la detección de movimiento es obtenida correctamente. Por lo tanto, una secuencia de imágenes puede ser modelado mediante aproximación por una gaussiana. Este modelo de substracción del fondo tiene un bajo coste computacional. Además, la convergencia del modelo se alcanza en un tiempo razonable y se adapta a los cambios lumínicos del ambiente, así como el ruido de adquisición por parte del sensor, véase Figura 6. Cabe destacar que el único reto en esta secuencia de imágenes se debe al ruido de adquisición del sensor. Por lo tanto, se espera que la substracción de fondo en un ambiente

exterior conlleve a un mayor reto debido a los cambios de iluminación, la sombra de objetos y los movimientos periódicos en el ambiente. Por último, se observa la mejora en la detección del movimiento al aplicar la operación morfológica de apertura, permitiendo descartar los efectos de ruido en la secuencia de imágenes.

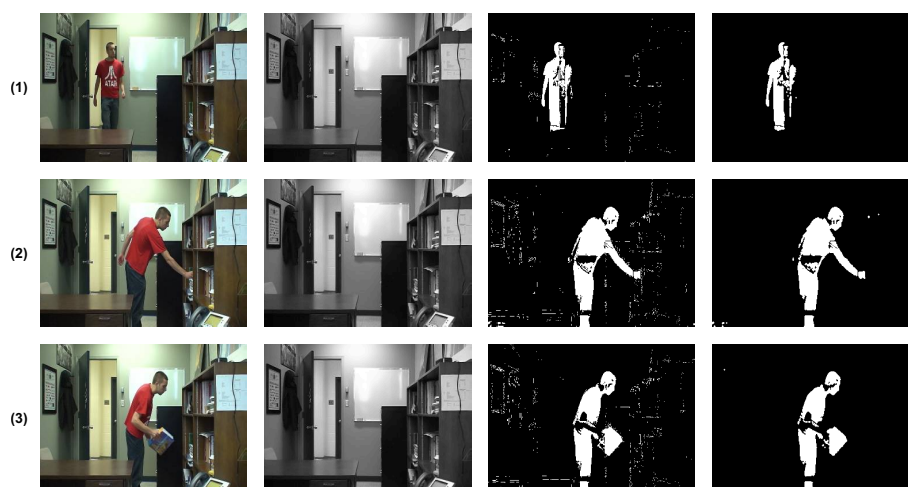


Figura 6. Substracción de fondo de una secuencia de imágenes en un escenario interior.

5 Conclusiones y Trabajos Futuros de Investigación

Este trabajo presenta una discusión de una de las tareas principales en el AA, la estimación de densidad. Primero, se revisaron los fundamentos históricos que llevaron al surgimiento del AA. Asimismo, se revisaron los paradigmas de aprendizaje y las tareas en el AA. Luego, se introdujo el concepto de estimación de densidad, así como, la topografía de enfoques y métodos que existen en la revisión de la literatura. Por otro lado, el objetivo de aplicar estos métodos para resolver diferentes problemáticas fue resuelto al presentar dos casos de uso; uno en el área de analítica de datos al realizar un análisis de frecuencias de los casos de COVID-19, y otro en el área de procesamiento de imágenes al desarrollar un modelo de substracción de fondo adaptable.

Cabe mencionar que la tarea de estimación de densidad como parte del paradigma de aprendizaje no supervisado, busca descubrir patrones o estructuras que representen información útil a cada problema a resolver. Por otro lado, la estimación de densidad es muchas veces una etapa previa para otras tareas de AA. Por ejemplo, en el caso del clasificador Bayesiano de tipo continuo, en algunos tipos de modelos gráficos probabilísticos para determinar relaciones causa-efecto, entre otras tareas, tal como, regresión y/o agrupamiento.

Por otro lado, los principales trabajos a futuro se centran en tres aspectos: *i)* desarrollar modelos de densidad de tipo multivariados, especialmente aquellos que trabajan con distintos componentes, conocidos como modelos de copula; *ii)* aplicar técnicas de búsqueda estocásticas, conocidas como algoritmos genéticos, para una mayor cercanía al óptimo global en la estimación de los parámetros; y *iii)* en modelos de densidad de tipo monovariados multimodales encontrar un criterio que determine el número de componentes que ajustan de forma eficiente la FDP.

Agradecimientos

Los autores agradecen al Centro de Investigación e Innovación en Ciencias de la Computación y Tecnología Educativa (CIICCTE) de la Facultad de Informática de la UAQ por brindar el espacio para llevar a cabo este trabajo.

Referencias

- [1] Deisenroth, M. P., Faisal, A. A., & Ong, C. S., Mathematics for Machine Learning. Cambridge University Press, 2020.
- [2] Murphy, K. P., Machine Learning: A Probabilistic Perspective. London, England: MIT Press, 2021.

- [3] Liu, K., Hogan, W. R., & Crowley, R. S., Natural language processing methods and systems for biomedical ontology learning. *Journal of biomedical informatic*, 163-179, 2011.
- [4] Mitchell, T. M., *Machine Learning*. McGraw-Hill, 1997.
- [5] Pearl, J., *Bayesian networks: A model of self-activated memory for evidential reasoning*. Irvine, California: UCLA Computer Science Department Technical Report 850021, 1985.
- [6] Xu, X., Yan, Z., & Xu, S., Estimating wind speed probability distribution by diffusion-based kernel density method. *Electric Power Systems Research*, 28-37, 2015.
- [7] Gómez-Lázaro, E. B.-M.-G. (2016). Probability density function characterization for aggregated large-scale wind power based on weibull mixtures. *Energies*.
- [8] Borunda, M. R.-V.-R.-S.-E., Long-term estimation of wind power by probabilistic forecast using genetic programming. *Energies*, 2020.
- [9] Zhu, S. Y. (2019). Gaussian mixture model coupled recurrent neural networks for wind speed interval forecast. *Energy Conversion and Management*.
- [10] Liu, Y., Qin, H., Zhang, Z., Pei, S., Jiang, Z., F. Z., & Zhou, J. (2020). Probabilistic spatiotemporal wind speed forecasting based on a variational bayesian deep learning model. *Applied Energy*.
- [11] Jiang, Y., Huang, G., Yang, Q., Yan, Z., & Zhang, C. (2019). A novel probabilistic wind speed prediction approach using real time refined variational model decomposition and conditional kernel density estimation. *Energy Conversion and Management*, 758-773.
- [12] Wang, H., & Zou, B. (2020). Probabilistic computational model for correlated wind speed, solar irradiation, and load using bayesian network. *IEEE Access*, 653-663.
- [13] He, Y., & Li, H. (2018). Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. *Energy Conversion and Management*, 374-384.
- [14] Jiang, W., & Tang, H. (2020). Distribution line parameter estimation considering dynamic operating states with a probabilistic graphical model. *International Journal of Electrical Power & Energy Systems*, 106-133.
- [15] Wahbah, M., Alhussein, O., & El-Fouly, T. Z. (2018). Evaluation of parametric statistical models for wind speed probability density estimation. *IEEE Electrical Power and Energy Conference*, 1-6.
- [16] Canton, D., Perez, J., & Jimenez, H. (2021). Bayesian analysis of geodependencies in wind speed. *XVII International Engineering Congress*, 1-7.
- [17] Ge, F., Ju, Y. Q., & Lin, Y. (2018). Parameter estimation of a gaussian mixture model for wind power forecast error by riemann l-bfgs optimization. *IEEE Access*, 92-99.
- [18] Canton-Enriquez, D., Niembro-Ceceña, J. A., Mandujano, M. M., Alarcon, D., Guerrero, J. A., Garcia, I. G., ... & Gutierrez-Lopez, A., Application of probabilistic models for extreme values to the COVID-2019 epidemic daily dataset. *Data in Brief*, 40, 107783, 2022.
- [19] Jiménez-Hernández, H., González-Barbosa, J., & Garcia-Ramírez, T., Detecting Abnormal Vehicular Dynamics at Intersections Based on an Unsupervised Learning Approach and a Stochastic Model. *Sensors*, 10(8), 7576-7601, 2010.
- [20] Wang, Y., Jodoin, P. M., Porikli, F., Konrad, J., Benezeth, Y., & Ishwar, P., CDnet 2014: An expanded change detection benchmark dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 387-394), 2014.