



**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE
MONTERREY**

Inteligencia artificial avanzada para la ciencia de datos I (Gpo 103)

“Entregable Reto”

Jose Pablo Cobos Austria A01274631

Profesor:

Dr. José Antonio Cantoral Ceballos

Santiago de Querétaro, 12 de septiembre del 2022

Selección de datos utilizados

Los datos que se seleccionaron para el análisis de esta evidencia fueron del dataset que contiene la siguiente información:

El conjunto de datos contiene 9568 puntos de datos recopilados de una central eléctrica de ciclo combinado durante 6 años (2006-2011), cuando la planta estaba configurada para funcionar a plena carga.

Las características consisten en variables ambientales promedio por hora:

- Temperatura (T)
- Presión ambiental (AP)
- Humedad relativa (RH)
- Vacío de escape (V)

Que se utilizaron para poder predecir la producción de energía eléctrica neta por hora (EP) de la planta.

Investigación sobre normativas que protejan ese tipo de datos

Ya establecido Cuáles son los datos que se van a analizar durante el desarrollo de esta evidencia, lo siguiente fue Investigar cuáles son las normativas que llegan a proteger este tipo de datos. Y para eso lo primero que se realizó fue clasificar o decidir qué tipo de datos estamos usando, que en este caso es información de una empresa particular que produce energía, y esto es muy importante saberlo porque se manejan de forma especiales a veces los casos entre datos personales y datos de empresas.

Ya también con ese segundo punto listo, lo siguiente fue realizar la investigación de la normativa que cubre ese tipo de datos, la cuál es la ***NORMA Oficial Mexicana NOM-151-SCFI-2016***, dónde nos establece cuáles son los formatos que se deben seguir Cuáles son las variables que se deben tomar en cuenta, cómo es que se deben guardar los datos, quién es el responsable en caso de situaciones adversas, cómo se deben de estar vigilando, entre muchas otras cosas más. No obstante es lo más importante que podemos hallar de esto, y también habla sobre los convenios internacionales que se tienen con diferentes países, ya que los datos obtenidos se recopilaron de una base de datos de Estados Unidos, por lo que fue necesario saber cómo funcionaban.

Entonces tomando en cuenta este documento, y tras haberlo leído, podemos decir que se respetó de forma correcta tanto los tratados internacionales como a nivel nacional el manejo de los datos ya que la fuente de donde se obtuvieron que es: UCI, qué es una base de datos internacional, dónde varias empresas personas, etcétera, donan o prestan dataset que tengan con el fin de implementar los diferentes tipos de de estudios, siendo en este caso lo que lo aplicamos para ver lo de machine learning. Además de que esta página está respaldada por la Fundación Nacional de Ciencias, una institución de prestigio y muy confiable de Estados

Unidos. Y todo esto en conjunto nos da como resultado un buen manejo de datos que se obtuvieron de forma correcta y sin infringir alguna ley

Análisis de la herramienta

Ahora tras haber analizado el dataset, los siguientes evaluar la herramienta que se desarrolló, que en este caso fue un pequeño modelo de regresión lineal cuyo fin principal es la predicción de la cantidad de energía que puede producir una planta, haciendo uso de múltiples variables ambientales y toda esta información fue recopilada de un recopilada en un periodo de 6 años.

Entonces ya explicado qué es lo que hace la herramienta, entre algunos de los acuerdos o normas que se encontraron sobre el manejo y la producción de energía, encontramos los siguientes: La **ISO 9001**, **ISO 14001** y **ISO 50001**. Qué nos hablan sobre los sistemas de gestión medioambientales, qué se requiere, cómo es que pueden disminuir las empresas impacto en el ambiente, normativas de seguridad y de salubridad con el fin de no abusar de sus empleados o a las condiciones laborales con el fin de incrementar sus ganancias, además de que la última es una normativa internacional para los sistemas de gestión de energía, qué busca mejorar la eficiencia y la optimización de estos mismos.

Entonces tomando en cuenta todas las normativas anteriormente mencionadas, sumándole la primera que se contó en el punto anterior, esta herramienta está totalmente libre de sesgo ético, ya que simplemente es su función o su utilización es para bienes de investigación y educativos, ya que se usó como una análisis para poder ver cómo es que una variable era afectada por otras y entender cómo funcionaban modelo regresión lineal.

Y entonces, tras haber analizado todo de herramienta haber leído las normativas, etc, podemos decir quién de una manera en la que se podría utilizar esta herramienta con malicia, es que una empresa trata de encontrar cuál es la variable que afecta más y realizar los ajustes a sus sistemas de producción energética, con el fin de poder crear más energía y estos mismos cambios resulten el pacto ambiental mayor al que deberían estar respetando. De igual manera otro uso que se le podría dar con malicia sería, la de implementarlo con una herramienta de predicción para ver cuánta energía se genera en el futuro y de esa manera estimar si alzar los precios o bajarlos de forma poco ética y puedan abusar del cliente.

Referencias:

- s/a. (2016). *NORMA Oficial Mexicana NOM-151-SCFI-2016*. Gob.mx. Recuperado el 14/09/2022 de: https://www.dof.gob.mx/normasOficiales/6499/seeco11_C/seeco11_C.html
- UCI. (s/f). *Combined Cycle Power Plant Data Set*. UCI.com. Recuperado el 14/09/2022 de: <https://archive.ics.uci.edu/ml/datasets/combined+cycle+power+plant>

- NQA. (s/f). *NORMAS EN LA INDUSTRIA ENERGÉTICA*. *nqa.com*. Recuperado el 14/09/2022 de: <https://www.nqa.com/es-mx/certification/sectors/energy>