

PRIMERA PARTE DEL PROYECTO

Dairo Joan Rivas García
Responsable

Raúl Ramos Pollas
Docente

Modelos y Simulación de Sistemas
Materia



**UNIVERSIDAD
DE ANTIOQUIA**

1 8 0 3

Universidad de Antioquia

Facultad de ingeniería

Medellín 2023

1. Descripción del problema predictivo

La actual situación climática en el mundo ha dado una relevancia a las energías renovables como una gran ayuda al medio ambiente y una necesidad de reducir la utilización y dependencia de combustibles fósiles como principales generadores de energía, es por esto que, a través de nuestro análisis de un conjunto de datos con información mundial del consumo de energía, queremos construir un modelo capaz de predecir en teravatios-hora el total de consumo de energía eléctrica proveniente de fuentes renovables en un país específico para un año dado. Con esto queremos motivar a inversionistas a participar en más proyectos renovables y que se puedan tomar acciones con tiempo que puedan impulsar estas empresas generadoras de energías renovables a crecer más en cada uno de los países.

2. Información de DataSet

El Dataset a utilizar proviene de la página de kaggle que la encontrara en el siguiente [link](#). Con archivo con la extensión .csv llamado "World Energy Consumption". Contiene información del consumo de energía eléctrica de alrededor de 242 lugares geográficos (países, continentes y organizaciones), se obtendrá la información de 22 años desde 1998 hasta el 2020, las cuales son el tiempo donde contiene la información de interés.

Algunos de las variables más importantes que contiene el archivo son:

Columna	Variable	Descripción
2	country	Ubicación geográfica
3	year	Año de la observación
60	hydro_share_elec	Proporción del consumo eléctrico que proviene de la hidroenergía
67	low_carbon_share_elec	Proporción del consumo eléctrico que proviene de fuentes de bajas emisiones. Esto es la suma de electricidad de energías renovables y nuclear
74	low_carbon_elec_per_capita	Consumo per cápita de electricidad a partir de fuentes de bajas emisiones (kilovatios-hora)
93	other_renewables_share_elec	Proporción del consumo eléctrico que proviene de otras fuentes de energía renovable
99	population	Población total del país
100	primary_energy_consumption	Consumo de energía primaria, medido en teravatios-hora

101	renewables_elec_per_capita	Consumo per cápita de energía primaria a partir de fuentes renovables (kilovatios-hora)
105	renewables_cons_change_twh	Cambio anual en el consumo de energía renovable, medido en teravatios-hora
107	renewables_elec_per_capita	Consumo per cápita de electricidad a partir de fuentes renovables (kilovatios-hora)
108	solar_share_elec	Proporción del consumo eléctrico que proviene de la energía solar
109	solar_cons_change_pct	Cambio porcentual anual en el consumo de energía solar
110	solar_share_energy	Proporción del consumo de energía primaria que proviene de la energía solar
111	solar_cons_change_twh	Cambio anual en el consumo de energía solar, medido en teravatios-hora
112	solar_consumption	Consumo de energía primaria a partir de la energía solar, medido en teravatios-hora
115	gdp	Producto Interno Bruto (PIB) total real, ajustado a la inflación
116	wind_share_elec	Proporción del consumo eléctrico que proviene de la energía eólica
117	wind_cons_change_pct	Cambio porcentual anual en el consumo de energía eólica
118	wind_share_energy	Proporción del consumo de energía primaria que proviene de la energía eólica
119	wind_cons_change_twh	Cambio anual en el consumo de energía eólica, medido en teravatios-hora
120	wind_consumption	Consumo de energía primaria a partir de la energía eólica, medido en teravatios-hora

3. Métricas de desempeño

- a. **Machine learning:** la métrica a utilizar para evaluar el desempeño será Error Cuadrático Medio Logarítmico (MSLE), el cual es una métrica de error que se utiliza para medir la discrepancia entre los valores predichos por un modelo y los valores reales; se preocupa más por el porcentaje de error que por la magnitud absoluta del error.

El MSLE penaliza más los errores en las predicciones de valores bajos que en las predicciones de valores altos. En este caso es útil

ya que se realizó un análisis previo a los datos y las variables en los cuales más del 85% de las columnas tienen una alta varianza y en este caso nos interesa analizar o son más importantes para nosotros los errores relativos.

Es decir por ejemplo si el valor real son 100 kilovatios/hora y el modelo predice 110 entonces el error relativo es de 10% y es este el valor que se tomará en cuenta al calcular el MSLE, pero si el valor que predijo el modelo es de 90, el error relativo es de -10%, este valor porcentual es que se tomará en cuenta al calcular el MSLE. De este modo tenemos que:

$$MSLE = \frac{1}{n} \sum_{i=1}^n (\log(p_i + 1) - \log(a_i + 1))^2$$

Donde:

- n es el número total de observaciones o puntos de datos.
- p_i es la predicción del modelo para la observación i .
- a_i es el valor real para la observación i .
- \log es el logaritmo natural.

El término “+1” se añade dentro de la función logarítmica para evitar tomar el logaritmo de cero (lo cual no está definido).

b. Métrica de negocio:

- El área de investigación de la empresa. Los analistas identificaron que la precisión en las predicciones de consumo de energía eléctrica, deben estar dentro del 5% del valor real, para ser consideradas como un valor aceptable y creíble frente a los inversionistas.

4. Desempeño en producción: Lo que se esperaría de este modelo es poder obtener la predicción del consumo de energía eléctrica renovable para un país específico en el próximo año. Con esta información, se desea obtener una mejor comprensión de cómo está evolucionando la transición hacia las energías renovables en diferentes partes del mundo. Esta información podrá ser útil para los responsables de la toma de decisiones que buscan implementar políticas de energía renovable, así como para las empresas que buscan invertir en tecnologías de energía renovable. Además, podría ayudar a identificar qué países están liderando la transición hacia las energías renovables y cuáles están rezagados, lo que podría ser valioso para los esfuerzos globales para combatir el cambio climático.