

Inteligencia artificial y paneles solares bifaciales para la optimización del autoconsumo energético¹

Artificial Intelligence and bifacial solar panels for optimizing energy self-consumption

Marta Sierra Segovia

IES Margarita Salas, Majadahonda, Madrid

La actual crisis energética y el acentuado cambio climático hacen patente la necesidad de optimizar el consumo energético. Gracias a la masiva cantidad de datos al alcance de todos y a los avances tecnológicos que facilitan su análisis, entre ellos el *deep learning* o aprendizaje profundo, es posible dar con soluciones innovadoras que abordan dichos problemas, entre otros, desde una nueva perspectiva. En el presente proyecto se ha desarrollado un modelo de análisis predictivo basado en redes neuronales recurrentes de memoria larga a corto plazo (LSTM). Dicho modelo tiene la capacidad de estimar la producción energética de una instalación de autoconsumo formada por paneles solares bifaciales en las futuras veinticuatro horas, utilizando los datos de los tres días anteriores. Para su desarrollo se procesaron un total de 61.320 datos de potencia obtenidos a partir de datos meteorológicos, valores de albedo correspondientes a siete superficies distintas y el diseño digital de una instalación fotovoltaica. Los resultados corroboraron la predictibilidad de la potencia eléctrica, puesto que el algoritmo adquirió una notable precisión, con un MSE de 0,012 y un RMSE de 0,109. Esto permite la adecuación de la demanda a la disponibilidad, optimizando así el consumo energético.

Palabras clave: Inteligencia artificial, albedo, red neuronal recurrente LSTM, predicción, autoconsumo

The current energy crisis and ongoing climate change shed light on the need to optimize energy consumption. Thanks to the massive amount of data available and the technological breakthroughs that facilitate its analysis, such as deep learning methods, it is possible to discover innovative solutions to address these problems. In this project, a predictive analysis model based on long short-term memory (LSTM) recurrent neural networks has been developed. This model has the capacity to estimate the energy production of a self-consumption installation made up of bifacial solar panels in the future twenty-four hours, using the data of the three previous days. For its development, a total of 61,320 power data was processed, which was obtained from meteorological data, albedo values corresponding to seven different surfaces and the digital design of a photovoltaic installation. The results confirmed the predictability of electrical power since the algorithm achieved remarkable accuracy, with an MSE of 0.012 and an RMSE of 0.109. This enables the adjustment of demand to availability, thus optimizing energy consumption.

Keywords: Artificial intelligence, albedo, LSTM recurrent neural network, prediction, self-consumption

Introducción

El 25 de septiembre de 2015, los 193 estados miembros de Naciones Unidas aprobaron la Agenda 2030 para el Desarrollo Sostenible, un ambicioso plan de acción comprometido con el planeta y sus habitantes. Con el fin de garantizar el acceso a una energía asequible, fiable, sostenible y moderna para todos se consensuó el objetivo de desarrollo sostenible (ODS) número 7. Acontecimientos recientes como la pandemia global de la Covid-19 o la actual guerra en Ucrania han sido agravantes de la crisis energética que el mundo arrastra desde principios de siglo, dejando al descubierto un claro déficit energético y provocando una inflación en los precios de la energía. Esta carencia ha tenido un notable impacto social negativo, no solo sobre los casi 90 millones de personas en

Asia y África, que por fin habían conseguido acceder a electricidad y, sin embargo, a día de hoy, no pueden pagar sus necesidades energéticas básicas, sino también sobre la población que sí tiene acceso a estos recursos. Este proyecto es una apuesta por una sostenibilidad ambiental, económica y social. Mediante el uso de inteligencia artificial (IA), propone un sistema de predicción de la producción energética en una instalación para el autoconsumo compuesta por paneles solares fotovoltaicos bifaciales.

Hoy en día, el carbón es una de las principales causas del cambio climático antropogénico, ya que contribuye con casi la mitad del total de emisiones de CO₂. Junto con otros combustibles fósiles supone un peligro para la salud pública, siendo todavía 2400 millones de personas las que dependen de estos para uso doméstico básico. Las energías renovables y modernas, como la estudiada en

¹ Una versión de este artículo fue seleccionado y participó impartiendo una ponencia en URANIA III, el Congreso Navarro de Investigación Joven, celebrado los días 11, 12 y 13 de junio de 2023.

este proyecto, la solar fotovoltaica, son una alternativa sostenible que permiten satisfacer las necesidades básicas sin suponer una amenaza para el medio ambiente o la salud de las personas vulnerables.

La actual crisis energética hace evidente tanto la necesidad de priorizar las energías sostenibles en el ámbito de la oferta energética, como la urgencia de reducir la demanda a través de la optimización de los recursos energéticos disponibles. La iniciativa de este proyecto facilita dicha optimización ya que permite al consumidor modificar sus hábitos de consumo en el hogar adaptándose a la producción energética. A pesar de la intermitencia de las condiciones climatológicas, la IA permite conocer por anticipado la disponibilidad de energía fotovoltaica. Esto supone una revolución en el ámbito de las renovables, ya que hace posible adecuar nuestra demanda a su disponibilidad. Una ventaja fundamental de la energía solar fotovoltaica, en especial si se usan paneles solares bifaciales, es su bajo coste de producción en comparación con otras energías. Además, la producción para el autoconsumo elimina los costes de transporte y las emisiones derivadas de este proceso. Esta mayor asequibilidad y la fiabilidad mencionada anteriormente hacen que la energía solar se pueda tomar en consideración para abastecer a los 733 millones de personas que no tienen acceso a electricidad.

Marco teórico

Energía solar fotovoltaica

La energía solar fotovoltaica se obtiene al transformar directamente la energía del sol en electricidad usando tecnologías basadas en el efecto fotoeléctrico (*¿Cómo funcionan las plantas fotovoltaicas?*, s. f.). Esta tecnología energética está en un constante desarrollo fomentado por las numerosas investigaciones vigentes, siendo una de las más prometedoras en vistas al futuro gracias a la reducción de costes que supone en la producción y el consumo energético.

La energía solar fotovoltaica está ligada al ODS 7: «Garantizar el acceso a una energía asequible, fiable, sostenible y moderna para todos» (Naciones Unidas, 2020). Algunas de sus ventajas:

- Es renovable y no contaminante, ya que no emite gases de efecto invernadero, genera residuos o consume combustibles fósiles.
- Está disponible en zonas aisladas donde hay acceso limitado a la electricidad.
- Impulsa la economía local al generar empleo y ofrecer electricidad sin tener que ser importada.
- Es una energía modular, es decir, se puede emplear a pequeña (generadores con autoconsumo) o gran escala (centrales fotoeléctricas).
- Es completamente competitiva con las energías convencionales ya que ha experimentado una drástica reducción de costes.

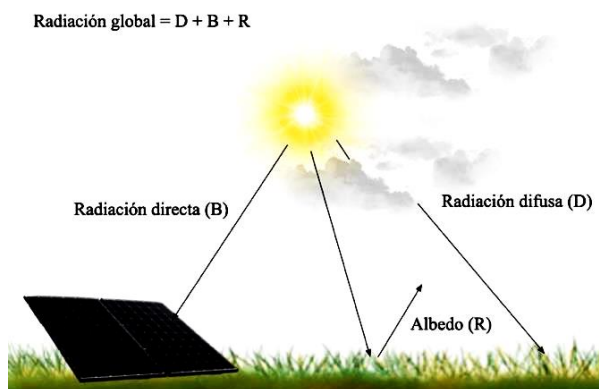
En cuanto a inconvenientes, destacan los problemas de almacenamiento y la intermitencia de la climatología. En cambio, se están produciendo avances en el ámbito de las baterías y la predicción de condiciones meteorológicas (Endesa, 2022).

Radiación solar

La radiación solar tiene su origen en el núcleo del sol y viaja a través del espacio en forma de ondas electromagnéticas hasta llegar a la atmósfera, donde sufre un proceso de atenuación al interactuar con las partículas que se encuentran en esta (IDEAM, s. f.). Esta atenuación puede darse por dispersión, absorción o reflexión. Una vez ha traspasado la atmósfera, llega a la superficie terrestre y es absorbida o reflejada. Como puede verse en la Figura 1, la radiación solar global está conformada por varios tipos de radiación solar:

- Radiación solar directa (B). Atraviesa la atmósfera y llega a la superficie sin haber sido dispersada.
- Radiación solar difusa (D). Llega a la superficie después de ser dispersada y cambiar de dirección.
- Radiación solar reflejada (R), también conocida como albedo, se define como el porcentaje de radiación solar reflejada por la superficie terrestre debido a un fenómeno denominado efecto albedo. Según la superficie en cuestión, el albedo toma valores distintos comprendidos entre 0 y 1, aunque su valor promedio es 0,3 o 30 % (Iberdrola, 2021).

Figura 1. Radiación solar.



Fuente. Elaboración propia.

La cantidad de radiación que incide en la superficie terrestre depende de la nubosidad y otras condiciones meteorológicas, la localización geográfica, la elevación del sol en el horizonte, y factores astronómicos como la época del año y las horas de luz (IDEAM, s. f.). Para medirla se dos magnitudes: la irradiancia (potencia de la radiación solar que incide en un instante sobre una superficie, W/m^2), o la irradiación solar (la energía de la radiación solar que alcanza a una superficie en un periodo de tiempo, J/m^2 o Wh/m^2) (HelioEsfera, 2020).

Para transformar la radiación solar en electricidad se utilizan paneles o módulos solares fotovoltaicos (Figura 2). Su origen se remonta a 1839, cuando el físico francés Alexandre-Edmond Becquerel descubrió el efecto fotovoltaico, aunque no fue hasta 1883 cuando el inventor estadounidense Charles Fritts fabricó la primera célula fotovoltaica (Ojeda-Durán, 2018). Inicialmente, su uso era aeroespacial, utilizándose para abastecer a los satélites de electricidad (Acciona, s.f.). En la actualidad, se emplean para suministrar electricidad a empresas, industrias, viviendas, zonas de difícil acceso al suministro eléctrico, etc. (BBVA, 2022).

Figura 2. Instalación de paneles solares

Fuente. Elaboración propia

Paneles solares

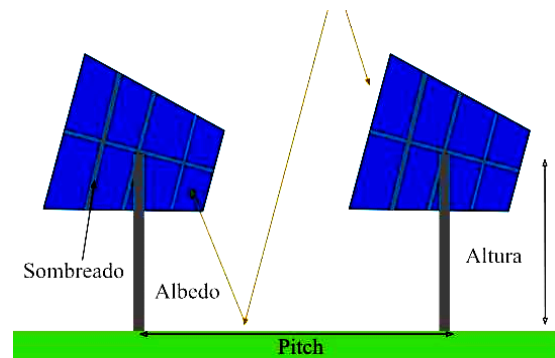
Los paneles solares fotovoltaicos actuales tienen una larga vida útil, de entre 25 y 30 años, lo cual hace que sean económicamente rentables, especialmente a largo plazo (Endesa, 2022). Están normalmente formados por 60 células fotovoltaicas, las cuales suelen ser de silicio (BBVA, 2022). Los de células de silicio monocristalino son los más caros y eficientes, mientras que los policristalinos y de capa fina están fabricados con materiales más baratos, tienen una eficiencia menor y son bastante demandados para el uso doméstico (Endesa, 2022), (Planas, 2020). Aparte de células fotovoltaicas, tienen otros componentes que les dan firmeza y operatividad, y protegen a las células (Planas, 2022). Algunos cuentan con un seguidor que cambia la posición del panel para maximizar su eficiencia. Por otro lado, elementos imprescindibles en las instalaciones fotovoltaicas son el inversor y el transformador.

Una característica esencial de los paneles fotovoltaicos es su eficiencia, referida al porcentaje de energía solar incidente en el panel que este es capaz de convertir en energía eléctrica. Esta depende principalmente del diseño y material del módulo. El rendimiento, en cambio, determina la energía eléctrica que se obtiene a partir de energía solar en unas condiciones específicas. Este depende de factores como la orientación e inclinación del panel. Cuanto más perpendiculares le lleguen los rayos del sol, mejor rendimiento tendrá el panel. Mientras que la orientación óptima viene determinada por el hemisferio en el que esté ubicado (en el hemisferio norte conviene que las placas estén orientadas hacia el sur, mientras que en el hemisferio sur conviene que estén orientadas hacia el norte), la inclinación viene dada por la latitud. En España, la inclinación óptima oscila entre los 20° y 40°. Otro factor que influye en el rendimiento es la temperatura de trabajo a la que está sometido el módulo, siendo aproximadamente 25°C la idónea (Planas, 2021), (Díaz, 2022).

Uno de los avances más prometedores de la tecnología fotovoltaica es el módulo bifacial. A diferencia de los paneles solares convencionales, monofaciales, los bifaciales producen electricidad a través de ambas caras al tener células fotovoltaicas en su cara posterior. Consecuentemente, mientras que la cara superior aprovecha la radiación directa y la difusa, la cara inferior aprovecha la reflejada o albedo. De este modo se incrementa la producción

de electricidad (Cambio energético, 2019). Como puede verse en la Figura 3, los factores que influyen en este aumento son:

- **Albedo.** Es el porcentaje de radiación solar que refleja una superficie, por lo que depende de la superficie en cuestión. En superficies claras, brillantes y de textura suave el albedo es alto, mientras que en superficies oscuras y rugosas es bajo. Cuanto más alto sea el albedo, más radiación solar llegará a la parte trasera del panel y más electricidad producirá.
- **Distancia entre paneles o *pitch*.** Cuanto más alejados estén los paneles, más electricidad producen. Se debe a que, al estar separados, la superficie reflectante útil de cada panel es mayor.
- **Altura.** Al elevar un panel bifacial a cierta altura, su temperatura desciende y su sombra se vuelve más clara y difusa, resultando en un aumento de la radiación reflejada recibida.
- **Sombreado.** La sombra producida por elementos de la propia estructura del panel como los cables y la fijación causan que el panel reciba menos radiación solar (Cambio energético, 2019).

Figura 3. Factores que influyen en el rendimiento de los paneles solares bifaciales

Fuente: Elaboración propia

Para incrementar el rendimiento del panel se puede utilizar un seguidor (BBVA, 2022). Según el Instituto de Investigación de Energía Solar de Singapur, la producción energética de los módulos bifaciales aumenta un 35% al ser combinados con seguidores de un solo eje, y un 40% al combinarlos con aquellos de dos ejes (Interempresas, 2021). Además, esta combinación permite generar la electricidad más barata de la historia (Energías Renovables, 2021). A pesar de todas las ventajas que presentan estos paneles, hay que considerar ciertos inconvenientes derivados de la heterogeneidad de la irradiancia que reciben las celdas en la cara posterior, como un aumento de las pérdidas de acoplamiento o los posibles puntos calientes. Esto conlleva una reducción de la eficiencia y potencia de salida, y la acelerada degradación de dicho módulo. Por último, como la corriente de cortocircuito de un *array* fotovoltaico es superior cuando la ganancia bifacial es mayor, se requiere de un inversor capaz de soportar la elevada corriente de cortocircuito de la instalación (Pérez, 2021). Estos paneles fabricados por primera vez en 1987 y se empezaron a comercializar en 1981, aunque su auge

llegó hace unos 6 o 7 años. El generador fotovoltaico bifacial más antiguo del mundo, mostrado en la Figura 4, se encuentra en el Instituto de Energía Solar de la Universidad Politécnica de Madrid (IES-UPM) (Higueras Rabadán, s.f.). En 2018, los paneles bifaciales representaban el 10 % de las instalaciones fotovoltaicas nuevas, y se espera que a lo largo de la década el porcentaje aumente a un 40 % (Interempresas, 2021).

Figura 4. Director del IES- UPM Carlos del Cañizo y el generador fotovoltaico más antiguo del mundo.



Fuente. Elaboración propia

Big Data o Macrodatos

Durante las últimas décadas, el volumen de datos digitales mundiales ha experimentado un crecimiento exponencial, pasando de 2 zettabits en 2010, a 64 en el año 2020 (Roa, 2021). Big Data es un término que hace referencia a un volumen masivo de datos y a su tratamiento y análisis. Como se muestra en la Figura 5, el tratamiento de datos sigue una estructura (Universidad Autónoma de Madrid, 2021).

Figura 5. Proceso de tratamiento de datos



Fuente. Elaboración propia

Cinco de sus características más destacables son:

- Volumen. La gran cantidad de datos disponibles que se necesita tratar.
- Variedad. Los datos son muy diversos ya que pueden tener diferentes formatos. Estos se dividen en tres tipos según su estructura: datos estructurados, semi-estructurados, y no estructurados.
- Velocidad. Los datos deben ser procesados casi en tiempo real.
- Veracidad. Los datos tienen que ser fiables para obtener información de calidad.
- Valor. Al analizar los datos se busca obtener información útil (Universidad Complutense de Madrid, 2020), (Lillo, 2021).

En lo que respecta al ODS 7, trabajar con tecnologías Big Data puede suponer una revolución en el sector

energético y un significativo avance hacia el objetivo (Naciones Unidas, s. f.).

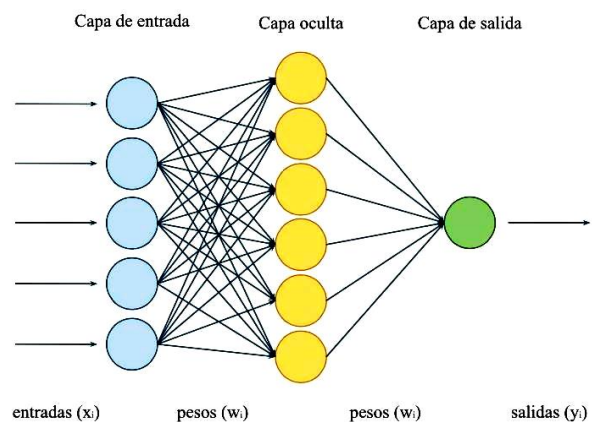
Debido al aumento de la demanda energética causado por el crecimiento demográfico y económico, y la escasez de combustibles fósiles, es necesario diseñar un nuevo modelo energético. Mediante el uso de Big Data se puede analizar todo el proceso relacionado con la energía y consecuentemente mejorar la eficiencia de su uso, garantizar el cubrimiento de demanda, hacer más sostenible el proceso, facilitar el uso extensivo de energía renovable, y reducir la huella de CO₂, la dependencia energética y los costes (Pequerul Herrero, 2017).

Inteligencia artificial

La inteligencia artificial (IA) es una técnica que pretende imitar la inteligencia humana a través de sistemas que aprenden mediante el análisis de datos. El puente entre la IA y el Big Data es el aprendizaje automático (*machine learning*), que permite la identificación de patrones en los datos mediante algoritmos entrenados para extraer conclusiones. Un subconjunto de este es el aprendizaje profundo (*deep learning*), basado en redes neuronales artificiales, las cuales imitan el funcionamiento de las neuronas humanas (Instituto de Ingeniería del Conocimiento, 2021).

Como se muestra en la Figura 6, las redes neuronales están formadas por capas de entrada, salida y ocultas, que procesan la información siguiendo un orden. Estas a su vez están formadas por neuronas o unidades que realizan una serie de operaciones matemáticas en los datos, y se conectan con todas las neuronas de capas anteriores y posteriores mediante pesos sinápticos. Dichos pesos regulan la información que se transmite a la siguiente capa y establecen la influencia de los valores de entrada para determinar el valor de salida (Amat Rodrigo, 2021).

Figura 6. Arquitectura de una red neuronal artificial



Fuente. Elaboración propia

Dentro de cada neurona tienen lugar dos operaciones:

- La suma ponderada de las entradas, multiplicando la suma de los valores de entrada, ponderados por sus pesos asociados, y sumándole el sesgo o *bias*. Se expresión matemática es:

$$f(x) = b + \sum_{i=1}^n w_i x_i$$

- Una función de activación que transforma los valores de entrada en valores de salida pertenecientes a un rango determinado y controla el flujo de información a la siguiente capa. En este proyecto se ha empleado sigmoide, la cual transforma los valores al rango (0,1). Su expresión matemática es:

$$f(x) = \frac{1}{1+e^{-x}}$$

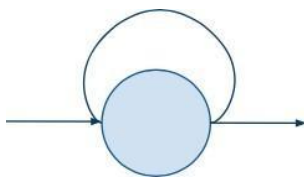
Otra función que toma un papel fundamental en el campo de las ANN es la función de pérdida, encargada de cuantificar la diferencia entre los valores predichos y los reales, es decir, dar un valor al error de predicción de la red, siendo 0 el valor mínimo. Para los problemas de regresión, la función de pérdida más utilizada es el error cuadrático medio (MSE, por sus siglas en inglés). Este calcula el error en un único punto elevando al cuadrado la diferencia entre el valor predicho y el estimado. Para obtener el error en el conjunto total de datos, se hace la media de los errores en todos los puntos. Su expresión matemática es:

$$L(w, b) = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

Para el entrenamiento de una ANN es necesario el pre-procesado de los datos, proceso en el que destaca la estandarización o escalado de estos con el fin de trabajar con datos que estén en la misma escala y evitar que independientemente de la relevancia de los datos, los de mayor escala tengan una mayor influencia en las predicciones. En cuanto al entrenamiento de la red, este consiste en el ajuste de los valores para los vectores de pesos con el fin de identificar las variables que más influyen en los resultados (Amat Rodrigo, 2021).

Existen varios tipos de ANN, cada uno conveniente para un tipo de problemas. Por ejemplo, cuando se trabaja con series temporales, las redes neuronales tradicionales no consiguen resultados fiables debido a que su arquitectura no les permite recordar valores anteriores. En cambio, las redes neuronales recurrentes (RNN), un tipo de ANN, tienen *loops* en su arquitectura que les permite recordar información de pasos previos. Estos *loops*, representados en la Figura 7, permiten que las neuronas reciban información de capas anteriores, además de sus propios valores de salida del instante de tiempo anterior (Torres, 2019).

Figura 7. Neurona de una RNN



Fuente. Elaboración propia

En cambio, cuando se trabaja con grandes cantidades de datos, se crea un *gap* temporal entre la información de tal manera que las RNN tampoco son capaces de almacenarla y recordarla. Las redes neuronales de memoria larga a corto plazo o LSTM (*long short-term memory*) son una evolución de las RNN, con una arquitectura más compleja. Tienen la capacidad de recordar información durante extensos intervalos de tiempo gracias a múltiples

celdas de memoria que se modifican de manera regulada a través de capas con funciones de pérdida, como la capa sigmoide o la tangente hiperbólica, que determinan qué información se debe olvidar, añadir, o recordar (González, 2020). Gracias a sus propiedades ya mencionadas, su uso se ha expandido en los ámbitos de predicción de series temporales, el reconocimiento de idiomas, etc. (Microsoft, 2022).

Objetivos e hipótesis

El objetivo principal del presente proyecto es:

- Programar un algoritmo basado en redes neuronales recurrentes de memoria larga a corto plazo o LSTM (*Long Short-Term Memory*), que sea capaz de realizar una predicción con precisión de la energía generada por una instalación fotovoltaica de autoconsumo, de placas bifaciales, en función del albedo de las superficies, en las veinticuatro horas posteriores

Como objetivos secundarios se contemplan los siguientes:

- Contribuir a la optimización del consumo energético ligado al ODS 7.
- Estudiar la producción de potencia eléctrica en módulos solares bifaciales.
- Investigar técnicas de IA para su aplicación en análisis predictivo.

Como hipótesis, se plantea que la inteligencia artificial permite ajustar el autoconsumo a la producción de energía.

Metodología

La metodología procedimental de este proyecto se basa en la predicción a corto plazo de la producción energética de los paneles solares bifaciales, en función de la potencia generada en las 72 horas precedentes al momento en cuestión, mediante el uso de un algoritmo basado en una red neuronal recurrente LSTM. El código de programación diseñado se encuentra en el Anexo I.

Herramientas empleadas

- *System Advisor Model (SAM)*. Un software gratuito que realiza predicciones de rendimiento y costes para proyectos relacionados con la industria de energías renovables (Freeman, s. f.).
- *JRC Photovoltaic Geographical Information (PVGIS)*. Una herramienta de la Unión Europea que cuenta con información meteorológica y de rendimiento energético. Utiliza imágenes de satélite y ofrece datos de toda Europa y África, y gran parte de Asia y América (EU ScienceHub, s. f.).
- *Jupyter Notebook* y *Anaconda*. Para desarrollar todo el código de programación se empleó *Jupyter Notebook*, un entorno de trabajo interactivo de código abierto. Es compatible con varios lenguajes de programación, incluyendo *Python*, el cual ha sido el utilizado en la elaboración del programa. Además, se puede utilizar online o en local (Santos, 2021). En el caso de este proyecto se accedió a él de forma local, a través de la instalación de la aplicación *Anaconda*.

- edX. Una plataforma de cursos en línea. Con el fin de adquirir los conocimientos necesarios para la elaboración del código de programación, se estudiaron los cursos “¡Conceptos básicos de Python para Data Science!”, “Analizando datos con Python” y “Visualizando datos con Python”, todos impartidos por la empresa IBM.

Procedimientos

Los pasos seguidos son los que aparecen en la Figura 8.

Figura 8. Procedimientos seguidos para la elaboración del programa



Fuente. Elaboración propia

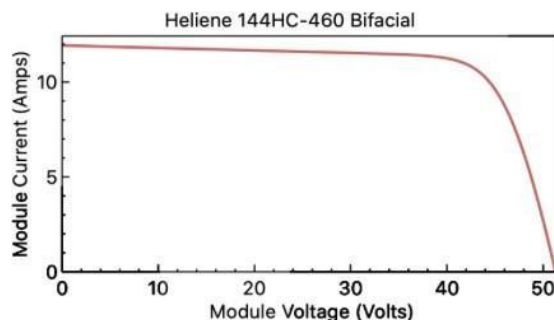
En primer lugar, se abordó la obtención de datos meteorológicos a través de PVGIS que más adelante se tuvieron que introducir en SAM. Como se muestra en la Figura 11, para descargarlos se seleccionó Majadahonda como localización y los datos del año típico meteorológico (TMY) como tipo de datos deseados, los cuales son un conjunto de datos meteorológicos representativos de una localización dada. Por último, se descendieron los datos en formato epw (*EnergyPlus Weather Data File*), un tipo de archivo compatible con SAM.

Una vez obtenidos estos datos, se procedió a descargar los datos de potencia en SAM. Esta aplicación se basa en el diseño personalizado de un modelo de rendimiento mediante la selección de los parámetros que mejor se ajusten a un sistema de autoconsumo alimentado por paneles bifaciales.

En primer lugar, se creó un nuevo modelo de rendimiento. Se seleccionó el modelo fotovoltaico detallado (*Detailed PV Model*), y se omitió el modelo financiero ya que no es necesario para predecir la potencia. En el modelo, los parámetros a ajustar se dividen en varios bloques:

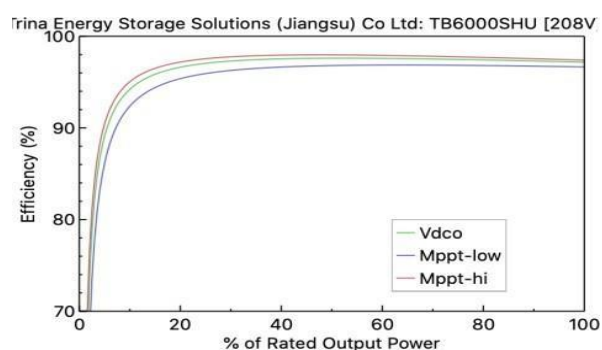
- *Location and resource.* Aquí se introdujeron los datos meteorológicos descendidos de PVGIS. Además, se pueden seleccionar distintos valores de albedo, aunque se dejó el coeficiente de albedo predeterminado, 0,2.
- *Module.* Aquí se puede elegir el módulo que formará parte del modelo. Se seleccionó un panel bifacial, el 'Heliene 144HC-460 Bifacial' compuesto de 72 células de silicio monocristalino y cuya área es de 2,210 m², tiene una eficiencia del 20.82% y una bifacialidad del 70% en condiciones de prueba estándar.

Figura 9. Curva intensidad-tensión del módulo seleccionado



Fuente: recuperado de SAM

Figura 10. Representación gráfica de la eficiencia del inversor



Fuente: recuperado de SAM

La curva intensidad-tensión de dicho módulo está representada en la Figura 9.

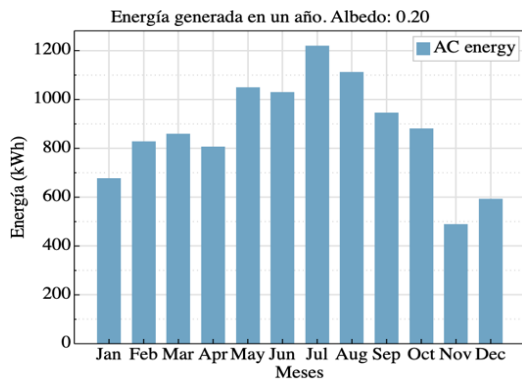
- *System Design.* Se puede elegir el número de módulos instalados por fila y columna. Como la instalación es para autoconsumo, se seleccionaron un total de 10 módulos, 2 filas compuestas por 5 módulos cada una. El área total del conjunto de módulos es de 22,100 m². En cuanto al sistema de seguimiento y a la orientación, se incorporó un seguidor de un solo eje y una inclinación de 35°.
- *Shading and Layout.* Se volvió a especificar el número de filas y columnas.
- *Inverter.* Seleccionar un inversor adecuado al modelo es esencial, puesto que este influye en gran medida en la potencia generada. Se seleccionó el inversor 'Trina Energy Storage Solutions (Jiangsu) Co Ltd: TB6000SHU [208V]', cuya eficiencia está representada en la Figura 14.

Los resultados obtenidos de la simulación del modelo aparecen en la Tabla 1 y la Figura 11.

Tabla 1. Características del modelo diseñado

Indicador	Valor
Energía CA anual	10485 kWh
Factor de capacidad de CC	26%
Rendimiento energético	2278 kWh/kW
Relación de rendimiento	0.84

Fuente. Elaboración propia

Figura 11. Energía anual (kWh) con albedo de 0,2

Fuente: Elaboración propia

Para poder trabajar posteriormente con los resultados, se descargaron en un archivo CSV, el cual incluye una tabla con valores horarios de un año. Aunque se pueden descargar distintos tipos de datos, se eligieron solamente los valores de la potencia generada en kW y el coeficiente de albedo.

Por último, se cambió el valor del albedo por valores conocidos asociados a distintas superficies sin modificar el resto de parámetros, y se descargaron los resultados obtenidos.

Para el procesamiento de datos, en primer lugar, se creó un nuevo cuaderno de Jupyter en donde desarrollar el código y se importaron una serie de librerías necesarias para el desarrollo del mismo. Las librerías son conjuntos de archivos de código y funcionalidades que facilitan la programación ya que reducen el tiempo de desarrollo y los costes (Datos.gob, 2022). Las que se emplearon para el desarrollo del programa son las siguientes:

- *Numpy*. Contiene funciones matemáticas avanzadas y proporciona una estructura de datos universal. Estas estructuras son matrices y vectores multidimensionales con una gran capacidad.
- *Pandas*. Una de las librerías más utilizadas en ciencia de datos. Esta librería ofrece estructuras de datos como las *Series* para conjuntos de datos unidimensionales y los *DataFrames* para conjuntos bidimensionales, y permite su manipulación a través de operaciones.
- *Matplotlib.pyplot*. *Matplotlib* es la librería estándar de *Python* y una de las más conocidas. Se emplea para generar múltiples tipos de gráficos o visualizaciones a partir de datos.
- *Seaborn*. Basada en *matplotlib*, permite la elaboración de gráficos estadísticos.
- *Scikit-learn*. Una librería usada para la construcción de modelos de aprendizaje automático y de análisis

de datos. Esta permite realizar aprendizajes supervisados y no supervisados.

- *Keras*. Es una librería de aprendizaje profundo muy útil para trabajar con redes neuronales ya que cuenta con un interfaz de alto nivel (Heras, 2020).

Se cargaron los datos de la potencia generada y el albedo en el cuaderno para poder trabajar con ellos. Se crearon 7 variables diferentes, una para cada conjunto de datos con un valor de albedo distinto, y se utilizó el método *pandas read_csv()* con la ubicación del archivo como único parámetro especificado. Dicho método lee un archivo csv (*comma-separated values*) y genera un *DataFrame*.

Se exploraron los datos para obtener las estadísticas y la forma del *DataFrame*, y asegurarse de que no hay valores nulos. Se empezó utilizando el método *.head()* para ver la estructura general de los *DataFrames* ya que este muestra sus cinco primeras filas. Además, se aplicó el método *.shape* para ver la forma del conjunto de datos. El resultado devuelto por este fue (8760, 3), lo que significa que el *DataFrame* tiene 3 columnas y 8760 filas. Mientras que las tres columnas corresponden al tiempo, la potencia generada y el albedo, las filas representan las 8760 horas de un año.

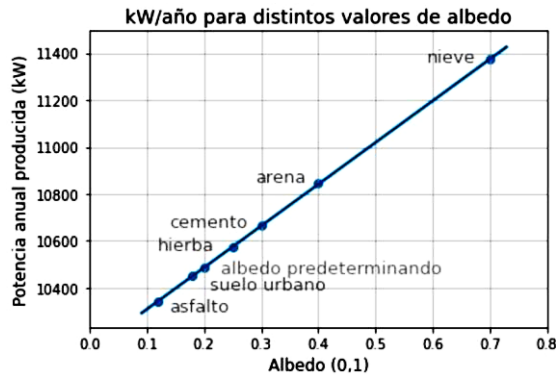
Para obtener las estadísticas de los conjuntos de datos y comprobar que no hay ningún valor atípico se utilizó el método *.describe()*, el cual devuelve una tabla con algunas estadísticas de las columnas formadas por valores numéricos. Por último, se empleó el método *.isnull().values.any()*, el cual devuelve un valor booleano, es decir, *True* o *False*, indicando si hay algún valor nulo, lo que podría tener una gran repercusión negativa en el análisis de datos. Como el *Output* fue *False*, se procedió con el preprocesado de datos.

Aunque no fuese necesario para el desarrollo del programa, se llevó a cabo un breve estudio sobre la correlación entre la potencia generada y el albedo. En primer lugar, se simplificaron los siete conjuntos de datos iniciales, creando una variable con la suma de los valores en la columna '*System power generated | (kW)*' para cada conjunto de datos, y otra variable con la media de dichos valores en cada conjunto. Una vez agrupados los datos, se crearon dos diccionarios. Uno formado por las claves '*Potencia anual producida (kW)*' y '*Albedo (0,1)*', y sus valores correspondientes, las siete variables creadas a partir de la suma de los valores del conjunto de datos inicial, y los siete valores de albedo que le corresponden a cada conjunto. El otro diccionario, en cambio, cuenta con las claves '*Media horaria potencia generada (kW)*' y '*Albedo (0,1)*', y contiene como valores las variables de la media y los valores de albedo. Después, ambos diccionarios se reestructuraron en un *DataFrame* de *pandas* para visualizarlos a través de un gráfico, empleando las librerías *seaborn* y *matplotlib.pyplot*, y obtener la correlación entre las dos variables.

Los gráficos resultantes fueron las Figuras 12 y 13, en las que se puede observar que existe una correlación lineal positiva entre la potencia generada y el albedo. Por último, se calculó numéricamente esta correlación mediante el uso del método *.corr()*, el cual usa el método de Pearson y devuelve un *array* de correlación. El coeficiente de correlación oscila entre -1 y +1, y es un indicador de cómo dos variables están relacionadas. En ambos casos, la potencia anual generada y la media horaria de la potencia generada, el coeficiente de correlación calculado es 0.999988. Como este coeficiente es muy próximo a uno,

su valor máximo, se puede deducir que la potencia depende en gran medida del albedo de la superficie en cuestión.

Figuras 12 y 13. Correlación entre la potencia anual generada y el albedo, y entre la media horaria de la potencia generada y el albedo.



Fuente. Elaboración propia

La reestructuración de datos se llevó a cabo para que estos adquiriesen el formato con el que la red neuronal LSTM puede trabajar. Para reducir los conjuntos de datos y las líneas de código necesarias, se agruparon los datos en dos conjuntos: entrenamiento y validación. Se decidió utilizar los datos calculados con el valor de albedo predeterminado (0,2) como datos de validación, y agrupar los otros seis conjuntos de datos para el entrenamiento de la red, aunque otras combinaciones también hubiesen sido válidas. De este modo, se redujo la cantidad de *DataFrames* de siete a dos.

A continuación, se modificaron las columnas de los *DataFrames*. Por un lado, se cambió el índice, puesto que en las series temporales de *pandas* es conveniente que el índice sea la columna del tiempo (*Time stamp*). Esto se realizó mediante el uso del método y los parámetros `.set_index('Time stamp', inplace=True)`, los cuales se aplicaron a los dos conjuntos de datos. Por otro lado, se eliminó la columna con los valores del albedo utilizando el método `.drop()`, puesto que la red neuronal solo necesita tomar como valores de entrada los datos de la potencia generada.

Otro paso importante a seguir es la normalización de los datos, lo cual consiste en reducir los datos a valores entre -1 y +1, o 0 y +1. Una de las escalas más comunes para normalizar datos es la escala Min Max. Esta convierte el valor mínimo del conjunto de datos en un 0 y el valor máximo en un 1, mientras que los valores intermedios toman valores decimales entre el 0 y el 1. La escala se representa como:

$$X' = (X - X_{min}) / (X_{max} - X_{min})$$

Para la normalización de los datos se empleó dicha escala importada de la librería *Scikit-learn*. Después de importar *MinMaxScaler*, se creó una variable que contuviera a la escala para poder utilizarla, y con el método

`.fit_transform()` de esta se ajustaron y transformaron los datos de entrenamiento y validación. Como este método devuelve un *array*, se volvieron a transformar los conjuntos de datos a *DataFrames*.

Una vez se obtuvieron los *DataFrames* con los datos escalados y solo los estrictamente necesarios, se estructuraron de tal manera que en una misma fila se encontrasen los 72 valores de la columna `'system power generated | kW'` anteriores al único valor anteriormente encontrado en la fila, y los 23 posteriores. De tal forma, en la fila correspondiente a una hora de cualquier día se encontraban, por orden cronológico, los valores de potencia generada en las 72 horas anteriores, el valor de la potencia generada en dicha hora y los valores de la potencia generada en las 23 horas posteriores. Esto se llevó a cabo mediante la definición de una función cuyos parámetros de entrada son los datos que se quieren modificar y el número de valores anteriores y posteriores a la hora en cuestión deseados. A estos dos últimos parámetros se les atribuyeron los valores predeterminados 72 y 24 respectivamente, lo que significa que, si no se especifica ningún valor para esos parámetros al ejecutar la función, esta utiliza los valores 72 y 24. La función devolvió otro *DataFrame* con el mismo número de filas que el tomado como parámetro, y su número de columnas multiplicado por 96, puesto que para cada columna del *DataFrame* original se generaron 72 columnas nuevas detrás y 23 delante. Al ejecutar la función con los datos de entrenamiento, el resultado fue un *DataFrame* con la forma (52560, 96), y el *DataFrame* obtenido al ejecutarla con los datos de validación tiene la forma (8760, 96).

Los *DataFrame* devueltos por la función tenían varios valores NaN. Esto se debe a que los primeros valores de los *DataFrames* no tenían suficientes valores precedentes, y los últimos no tenían suficientes valores siguientes. Para solucionar el problema se eliminaron todas las filas con valores NaN, es decir, las primeras 72 filas y las últimas 23 de cada *DataFrame*. Se empleó el método `.drop()` y se especificaron los intervalos de datos que se querían eliminar.

Se dividieron los datos en valores de entrada y de salida, es decir, valores que utiliza la red como *input* y valores que tiene que predecir. En primer lugar, se obtuvo un *array* con los valores de los *DataFrames*, para luego poder crear dos variables: 'X' e 'Y'. Al conjunto 'X' (*input*) se le asociaron los primeros 72 valores de cada columna del *DataFrame*, y al conjunto 'Y' los 24 últimos. Este proceso se llevó a cabo con ambos conjuntos de datos: los datos de entrenamiento y los datos de validación.

Por último, como los valores de entrada tenían dos dimensiones y las redes neuronales recurrentes LSTM funcionan con *inputs* en formato de tres dimensiones, se aplicó el método `.reshape()` a los valores de entrada de entrenamiento y validación para cambiarlos de forma. Los resultados fueron dos *arrays* multidimensionales, el de entrenamiento con la forma (52465, 1, 72), y el de validación con la forma (8665, 1, 72), siendo el valor intermedio el correspondiente a los pasos de tiempo.

Para el desarrollo y entrenamiento del modelo de análisis predictivo mediante el uso de una red neuronal LSTM en primer lugar se importaron las librerías necesarias para desarrollar el código, como las librerías de *deep learning* y la necesaria para calcular el error del programa.

A continuación se creó un modelo de *deep learning* de clase *Sequential* y se le añadieron varias capas con el método `.add()`. Primeramente, se agregó una capa de la red neuronal recurrente LSTM, cuyos parámetros son el número de neuronas y la forma del *array* de entrada, donde hay que establecer el número de pasos de tiempo y de indicadores. En cuanto a las neuronas, se seleccionó un

valor de 100, y el *array* de entrada se formó por las dos últimas variables de los *arrays* tridimensionales previamente generados, lo cual coincide con los pasos de tiempo (1) y su longitud (72).

La segunda capa que se añadió fue una capa de abandono o exclusión (*Dropout*). La función de esta capa es evitar el sobreajuste u *overfitting*, el fenómeno de sobreentrenar a un modelo de tal manera que funcione muy bien con los valores de entrenamiento, pero peor con los valores de validación. Esto causa que disminuya su porcentaje de acierto. La tasa de abandono elegida fue del 20 %, lo que significa que cada neurona tiene una probabilidad de permanecer desactivada del 20 %.

La última capa que se agregó fue una capa densa, también conocida como capa completamente conectada. En este tipo de capas, cada neurona está conectada con todas las neuronas de la capa anterior. El número de neuronas de esta capa concuerda con el número de valores de salida, el cual viene establecido por la segunda variable del *array* bidimensional de la forma del conjunto de datos de validación (24).

Una vez añadidas las capas al modelo, se procedió a compilarlo. Se seleccionó el error cuadrático medio (MSE) como función de pérdida, y *adam* como optimizador. Además, se incluyó en el código una parada temprana o *EarlyStopping*, una técnica de regularización que evita el sobreajuste.

Finalmente, se entrenó el modelo para que este sea capaz de realizar predicciones. En esta parte del código hubo que especificar varios parámetros:

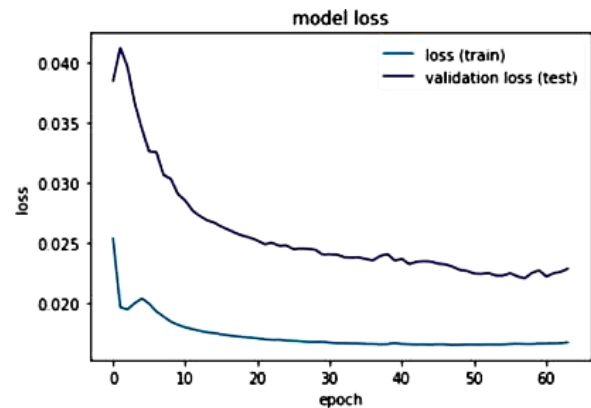
- Los valores de entrada y de salida con los que entrenar el modelo, en este caso los datos de entrenamiento divididos en valores de entrada y de salida.
- El número de ciclos o *epoch*, que son las veces que cada ejemplo pasa por la red. Este parámetro es bastante importante puesto que un número muy bajo de este causaría *underfitting* y uno muy elevado causaría *overfitting*. El valor seleccionado en este caso ha sido 100 ciclos.
- Devolución de llamada o *callbacks*, que incluye la parada temprana diseñada. En este caso detiene el entrenamiento si la pérdida de validación no disminuye durante seis ciclos consecutivos.
- El *batch size* o tamaño de lote, que determina la memoria del modelo ya que establece el número de ejemplos que pasan por el modelo en cada ciclo. Después de ejecutar el código repetidas veces probando distintos valores, se eligió el valor 85 puesto que con este se obtuvo el mínimo error.
- Los datos con los que validar el modelo, en este caso los datos de validación divididos en valores de entrada y de salida.
- La *verbose*, la cual simplemente indica el modo de verbosidad. No tiene ningún efecto en los resultados del modelo, pero indica cómo se muestra el resumen de cada ciclo al entrenar el modelo. Se eligió el valor dos, es decir, que cada ciclo sea resumido en una línea de código.
- *Shuffle*. Se puede elegir si los datos son desordenados automáticamente antes de entrenar el modelo. En el caso de las series temporales, este parámetro se tiene que igualar a *False* debido a que los datos siguen un orden cronológico que influye en la variable dependiente (González, 2020).

Resultados

Una vez entrenado el modelo, se representaron las curvas de aprendizaje, visualizaron los resultados y calculó el error del programa para comprobar que el modelo hubiese aprendido correctamente, que su funcionamiento fuese óptimo y, por tanto, que el modelo fuese fiable.

Se representaron gráficamente las curvas de aprendizaje de los ciclos frente a los valores de pérdida en los conjuntos de datos de entrenamiento (*loss*) y validación (*validation loss*). En la Figura 14 se puede ver cómo, aunque el número de ciclos establecido había sido 100, gracias a la parada temprana, el entrenamiento del modelo se detuvo tras el ciclo 64 para evitar un sobreajuste. Las pérdidas alcanzan valores bajos y estables, llegando al valor mínimo de pérdida de validación (0,0165) en el ciclo 58.

Figura 14. Curvas de aprendizaje del modelo. Proceso de entrenamiento 1.

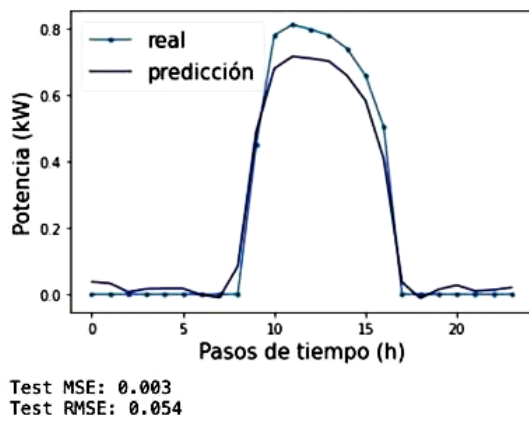


Fuente: Elaboración propia

El error cuadrático medio (MSE) calculado y su raíz (RMSE) comparando todos los valores de validación reales y los valores predichos correspondientes son 0,023 y 0,151. Como es un error bajo, se entiende que el programa tiene un buen funcionamiento, y se generaron gráficas para visualizar este error.

En primer lugar, se representó la predicción realizada frente a los valores reales de potencia generada durante el transcurso de un día. Por conveniencia, se eligió el primer día del conjunto de datos de validación. En la Figura 15 se puede ver cómo la predicción se aproxima a los valores reales y, en términos generales, el comportamiento del programa es correcto. Sin embargo, no llega a alcanzar los valores más altos. El MSE en este conjunto reducido de datos es 0,003 y el RMSE es 0,054.

Figura 15. Valores predichos y reales de potencia generada en un día. Proceso de entrenamiento 1.



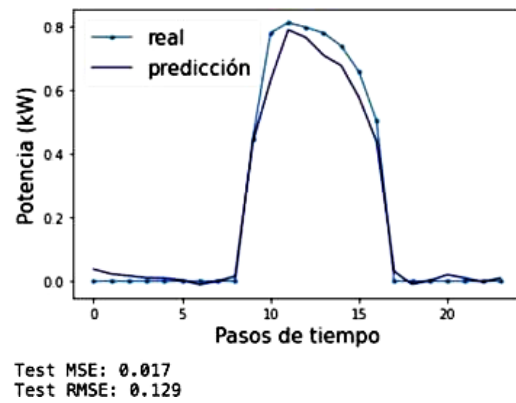
Fuente: Elaboración propia

A continuación, con el fin de evaluar el comportamiento progresivo del modelo, se llevó a cabo una comparación de los primeros valores predichos frente a los últimos. Suponiendo que se realiza una predicción cada hora durante un día, se representaron los primeros valores sugeridos en cada predicción comparados a los valores reales correspondientes, y en otro gráfico los últimos valores sugeridos. Las Figuras 16 y 17 muestran cómo el desempeño del programa empeora en la última hora predicha con respecto a la primera. Además, esta leve caída de rendimiento también se representa numéricamente, puesto que mientras que el MSE del primer valor de todas las predicciones es 0,019 y el RMSE es 0,129, cuando se considera el último valor predicho el error aumenta, siendo 0,022 el MSE y 0,147 el RMSE.

Con el fin de intentar conseguir una mejora en el rendimiento del programa, se llevaron a cabo una serie de modificaciones en el código y se volvió a ejecutar. Para hacer más complejo el código y aumentar la precisión del modelo se añadió una segunda capa de red neuronal LSTM con los mismos parámetros que la ya encontrada en el código, acompañada de otra capa de abandono. Además, se eliminó la parada temprana para ver la evolución de los valores de pérdida a lo largo de 100 ciclos.

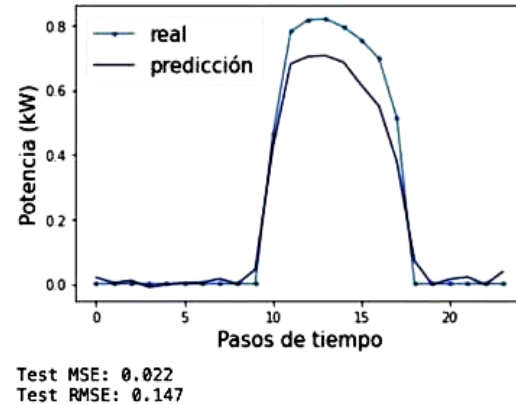
Los resultados obtenidos fueron significativamente mejores. El MSE descendió a 0,012, levemente superior a la mitad del error en el experimento anterior, y el RMSE a 0,109. La Figura 31 muestra cómo se alcanzan valores de pérdida menores, alcanzando los mejores resultados en el último ciclo, con un valor de pérdida de validación de 0,018. En torno al ciclo 10, el modelo experimenta un aumento en los valores de pérdida para posteriormente recuperar una tendencia descendente que perdura hasta el último ciclo. Si no se hubiese eliminado la parada temprana, el entrenamiento del modelo se hubiese detenido en torno al ciclo 20, lo que hubiese impedido obtener los resultados alcanzados sin este parámetro. Además, cabe destacar que, como se puede ver en la figura, las pérdidas de validación alcanzan valores inferiores a las pérdidas de entrenamiento, lo que indica un óptimo aprendizaje del modelo.

Figura 16 y 17. El primer valor de salida predicho frente al real, y el último valor de salida frente al real. Proceso de entrenamiento 1.



Fuente: Elaboración propia

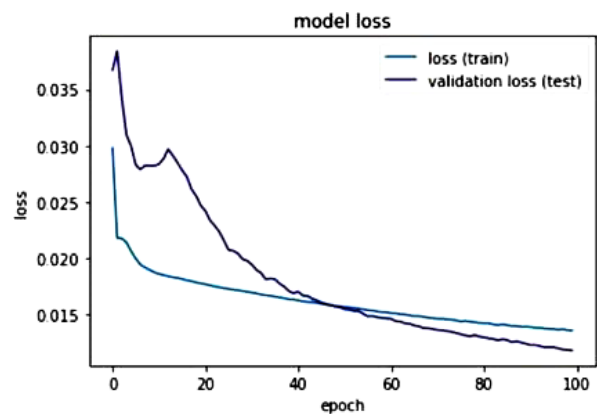
Figura 18. Curvas de aprendizaje del modelo. Proceso de entrenamiento 2.

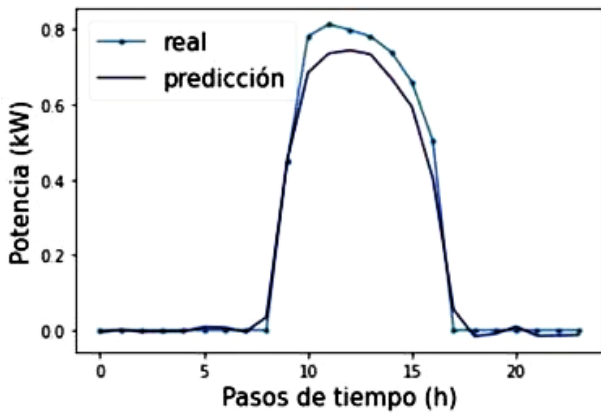


Fuente: Elaboración propia

Se volvió a representar la predicción del primer día ante los valores reales correspondientes en la Figura 19. En efecto, la predicción ha adquirido más precisión, lo que se puede ver gráficamente y a través del error numérico, ya que el MSE y en RMSE son menores que en el proceso de entrenamiento anterior.

Figura 19. Valores predichos y reales de potencia generada en un día. Proceso de entrenamiento 2.



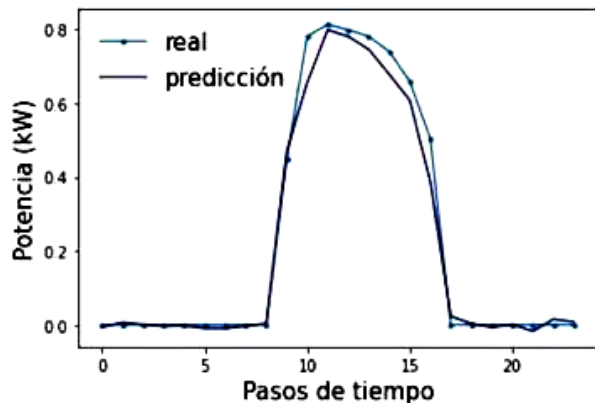


Test MSE: 0.002
Test RMSE: 0.043

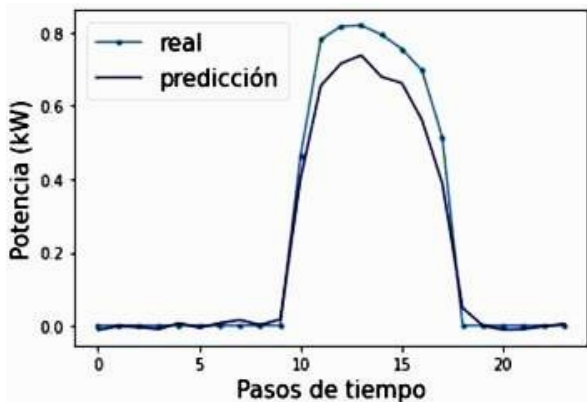
Fuente: Elaboración propia

Por último, se volvieron a representar los primeros y los últimos valores predichos. Como se puede ver en las Figuras 20 y 21, sigue existiendo una diferencia entre los dos conjuntos, siendo el primero el que presenta los mejores resultados. Aunque la diferencia del error entre ambos se haya mantenido, los dos han disminuido significativamente con respecto al proceso de entrenamiento anterior, y cabe destacar que la predicción de los primeros valores alcanza una aproximación bastante realista.

Figura 20 y 21. El primer valor de salida predicho frente al real, y el último valor de salida frente al real. Proceso de entrenamiento 2



Test MSE: 0.008
Test RMSE: 0.089



Test MSE: 0.014
Test RMSE: 0.118

Fuente: Elaboración propia

Conclusiones

A partir de los resultados obtenidos se pudo llegar a las siguientes conclusiones:

- La producción energética de una instalación de paneles solares bifacial para el autoconsumo se puede predecir con un grado de precisión elevado gracias a la inteligencia artificial, lo que se podría aplicar a otros tipos de paneles solares o a otras fuentes de energía renovables. El uso de un algoritmo basado en redes neuronales recurrentes LSTM ha permitido obtener un modelo predictivo con un MSE de 0,012 y un RMSE de 0,109.
- El modelo de predicción de energía fotovoltaica, gracias a su alto nivel de precisión, podría ser implementado en los hogares con el fin de optimizar el consumo energético.
- La energía solar fotovoltaica es una alternativa clave de los combustibles fósiles, no solo por su bajo coste y sostenibilidad, sino por la capacidad de predicción que permite la adecuación de la demanda a la disponibilidad.
- Los paneles solares bifaciales, instalados en superficies claras y lisas, es decir, con altos coeficientes de albedo, suponen un aumento de la producción y una disminución de costes.
- Existe una correlación lineal positiva entre la potencia generada por un panel solar bifacial y el albedo de la superficie en la que se encuentra el panel. El coeficiente de correlación entre ambas variables es, según lo calculado, 0,999988, lo que se aproxima al valor máximo que puede tomar dicho coeficiente, 1.
- Cuando se trabaja con redes neuronales artificiales, la elección de los hiperparámetros es clave para el buen funcionamiento del programa, y es necesario disponer de una amplia cantidad de datos para entrenar el modelo y obtener resultados óptimos.
- En el programa diseñado en este proyecto, el uso de dos capas de red neuronal LSTM en vez de una supone una disminución del error y un aumento de la precisión.

Algunas de las futuras líneas de investigación pueden suponer la adaptación del modelo de análisis predictivo basado en redes neuronales recurrentes LSTM para:

- Su implementación en países en vías de desarrollo, junto con un plan social y económico que permita garantizar el acceso a una energía moderna y sostenible, como la energía solar fotovoltaica, a la población más desfavorecida. Para ello será necesario adaptar el software a los datos disponibles en dichos entornos y a los dispositivos de los que dispone la población: por ejemplo, uso de móviles en lugar de ordenadores, etc.
- Aumentar la escala a un nivel comunitario o de ciudad, para buscar un mayor impacto de ahorro energético y una notable reducción del impacto medioambiental. Esto puede suponer, por ejemplo, realizar un software de predicción energética para que sea gestionado por una comunidad de vecinos o un municipio, acompañado de un sistema de información a los usuarios finales.

- Aplicar este tipo de modelo a otros sectores de alta demanda energética como puede ser el sector del transporte o el de la industria. Para ello será necesario abordar un software más profesional donde se puedan tener en cuenta otras variables y se pueda trabajar con sistemas de generación de energía de mayores dimensiones y que utilicen distintas fuentes primarias: además de fotovoltaica, eólica, solar térmica, geotérmica, etc.

Bibliografía

- Acciona. (s. f.). *Energía solar fotovoltaica y su contribución | ACCIONA | business as unusual*. Recuperado 20 de julio de 2022, de <https://bit.ly/3V2m70V>
- BBVA. (2022, 8 abril). *¿Qué son los paneles solares, cómo funcionan y cuál es su futuro?* Recuperado 27 de julio de 2022, de <https://bit.ly/3HLLZw3>
- Borja - Academia Energía Solar. (2021, 11 febrero). *Ten en cuenta ESTO si usas PANELES SOLARES BIFACIALES*[Video]. YouTube. Recuperado 19 de octubre de 2022, de <https://bit.ly/3hDnxbZ>
- Cambio energético. (2019, 1 agosto). *PLACAS SOLARES BIFACIALES*. Recuperado 9 de agosto de 2022, de <https://bit.ly/3Wnsgfh>
- Cambio energético. (2019, 26 febrero). *¿CUÁL ES LA ÚLTIMA TECNOLOGÍA EN MÓDULOS FOTOVOLTAICOS?* Recuperado 9 de agosto de 2022, de <https://bit.ly/3Yywjco>
- Datos.gob. (2022, 3 mayo). *11 librerías para crear visualizaciones de datos*. Recuperado 7 de septiembre de 2022, de <https://bit.ly/3G3cRws>
- Díaz, A. (2022, 22 junio). *Orientación e inclinación de las placas solares*. Otovo Blog. Recuperado 28 de julio de 2022, de <https://bit.ly/3G2U84i>
- Elvira Lillo. (2021, 16 septiembre). *Un mundo Big Data: origen y futuro del poder de los datos* [Video].
- Endesa. (2022, 30 junio). *¿Cómo se produce la energía solar?* Recuperado 20 de julio de 2022, de <https://bit.ly/3PQAsEh>
- Endesa. (2022, junio 29). *¿Cómo funcionan los paneles solares?* Recuperado 21 de julio de 2022, de <https://bit.ly/3uZLRl1>
- Energías Renovables. (2021, 20 julio). *El generador fotovoltaico bifacial más antiguo del mundo tiene 34 años, está en Madrid y sigue operativo*. Energías Renovables, el periodismo de las energías limpias. Recuperado 9 de agosto de 2022, de <https://bit.ly/3jbFtei>
- EU ScienceHub. (s. f.). *PVGIS Photovoltaic Geographical Information System*. Recuperado 10 de agosto de 2022, de <https://bit.ly/3PzOtpq>
- Freeman, J. (s. f.). *System Advisor Model (SAM)* [Diapositivas]. NREL. Recuperado de <https://bit.ly/3hyHoce>
- González, J. M. C. (2020, 23 junio). *Tuneando los hiperparámetros de una red neuronal LSTM para obtener un aprendizaje más eficiente*. LinkedIn. Recuperado 11 de septiembre de 2022, de <https://bit.ly/3j6oYjn>
- HelioEsfera. (2020, 19 agosto). *Irradiancia, irradiación y radiación solar*. <https://bit.ly/3W8Lok6> Ojeda-Durán, Esteban. (2018). *Historia, Desarrollo y Actualidad de las Celdas Solares*.
- Heras, J. M. (2020, 10 octubre). *15 Librerías de Python para Machine Learning*. lArtificial.net. Recuperado 7 de septiembre de 2022, de <https://bit.ly/3BNZEVV>
- Higuera Rabadán, E. (s.f.). *Estamos jugando en la 'primera división' de la investigación en energía solar*. Recuperado 9 de agosto de 2022, de <https://bit.ly/3G3PdQQ>
- Iberdrola (s. f.). *¿Cómo funcionan las plantas fotovoltaicas?* Recuperado 19 de julio de 2022, de <https://bit.ly/3hygiSF>
- Iberdrola. (2021, 22 abril). *Radiación solar: ¿cuál es su impacto sobre el planeta y el ser humano?* Recuperado 22 de julio de 2022, de <https://bit.ly/3FFZsJB>
- IDEAM. (s. f.). *Características de la radiación solar*. Recuperado 25 de julio de 2022, de <https://bit.ly/3Vgodk7>
- IDEAM. (s. f.). *La radiación solar y su paso por la atmósfera*. Recuperado 25 de julio de 2022, de <https://bit.ly/3hxA8xe>
- Interempresas. (2021, 28 octubre). *Tecnología bifacial, el futuro de la industria solar fotovoltaica*. Recuperado 9 de agosto de 2022, de <https://bit.ly/3Fxl7E>
- Machine Learning y Deep Learning* (2021, 21 enero). Instituto de Ingeniería del Conocimiento. Recuperado de <https://bit.ly/3BKz6N6>
- Microsoft. (2022, 29 noviembre). *Aprendizaje profundo frente a aprendizaje automático en Azure Machine Learning*. Microsoft Learn. Recuperado de <https://bit.ly/3WuZmLo>
- Muñoz Jimenez, A. (2013-2014). *Modelos de predicción a corto plazo de la generación eléctrica en instalaciones fotovoltaicas* [tesis doctoral]. Universidad de la Rioja.
- Naciones Unidas. (2020, 2 julio). *Objetivo 7: Garantizar el acceso a una energía asequible, segura, sostenible y moderna*. Objetivos de Desarrollo Sostenible. Recuperado 7 de junio de 2022, de <https://bit.ly/3BFKpy7>
- Naciones Unidas. (s. f.). *Macrodatos para el desarrollo sostenible*. Recuperado 11 de junio de 2022, de <https://bit.ly/3WkqXiI>
- Pequerul Herrero. (2017, 17 julio). *Big Data & Eficiencia Energética. Un nuevo modelo energético*. eSMARTCITY. Recuperado 4 de febrero de 2022, de <https://bit.ly/3jaO5Sg>
- Planas, O. (2020, 28 marzo). *Tipos de paneles fotovoltaicos: descripción y rendimiento*. Energía solar. Recuperado 27 de julio de 2022, de <https://bit.ly/3FYtR7h>
- Planas, O. (2021, 16 julio). *Eficiencia de los paneles solares fotovoltaicos*. Energía solar. Recuperado 28 de julio de 2022, de <https://bit.ly/3BKIYhZ>
- Planas, O. (2022, 20 junio). *Partes de un panel solar, componentes y estructura*. Energía solar. Recuperado 21 de julio de 2022, de <https://bit.ly/3FFZVvl>
- Redes neuronales con Python* by Joaquín Amat Rodrigo, available under a Attribution 4.0 International (CC BY 4.0) at <https://bit.ly/3YwzdOk>
- Roa, M. M. (2021, 22 octubre). *El Big Bang del Big Data*. Statista Infografías. Recuperado 11 de junio de 2022, de <https://bit.ly/3YuXAfB>
- Santos, P. R. D. L. (2021, 16 junio). *Python para todos (2): ¿Qué son los Jupyter Notebooks?* Telefónica, Think Big. Recuperado 7 de septiembre de 2022, de <https://bit.ly/3HHyeF2>
- Torres, J. (2019, 22 septiembre). *Redes Neuronales Recurrentes*. Jordi TORRES.AI. Recuperado 1 de octubre de 2022, de <https://bit.ly/3FAiLnr>
- Universidad Autónoma de Madrid. (2021, 14 julio). *Big Data*. Instituto de Ingeniería del Conocimiento. Recuperado 7 de junio de 2022, de <https://bit.ly/3hwE2qn>
- Universidad Complutense de Madrid. (2020, 17 septiembre). *¿Qué es Big Data?* Master Big Data. Recuperado 11 de junio de 2022, de <https://bit.ly/3VbjouW>

Fecha de recepción: 15 de octubre de 2022
Fecha de aceptación (provisional): 11 de marzo de 2023
Fecha de aceptación (definitiva): 16 de junio de 2023



Copyright ©
Asociación por la Investigación
en Educación Secundaria
AINVES

C/ Carlos II, 17
Madrid
28022