

‘Big data’ para comprender el entorno: estudio de la transición a vehículos eléctricos

Understanding the environment with “Big Data”: a study of the transition to electric vehicles

Asier Murcia Gila¹ y Manuel Enrique Islán Marcos²

Resumen

En el presente trabajo se expone de forma práctica el big data o análisis masivo de datos como herramienta para el estudio del entorno que rodea al diseño industrial. Se presenta su estructura general, así como sus nuevas técnicas de procesamiento de datos, como redes neuronales. Se propone un análisis de la transición de vehículos de combustión a vehículos eléctricos en España, al ser un tema de gran incertidumbre e interés industrial. Se hace uso de datos históricos y actuales recopilados de distintas fuentes para la construcción de un modelo predictivo de las tendencias de venta, evolución del parque vehicular y del impacto energético que supondrían los vehículos eléctricos en la red.

Palabras clave

Diseño industrial, big data, aprendizaje automático, redes neuronales, vehículos eléctricos, BEV.

Abstract

In this paper Big Data or massive data analysis is exposed in a practical view as a tool for studying the industrial design environment. Its general structure is presented, as well as its new data processing techniques, such as neural networks. An analysis of the Spanish transition from combustion vehicles to electric vehicles is proposed, since it is uncertain and has industrial interest. Historical and current data collected from different sources are used to build a predictive model of sales trends, the evolution of the vehicle fleet, and the energy impact that electric vehicles would have on the network.

Keywords

Industrial design, Big Data, machine learning, neuronal networks, electric vehicles, BEV.

Recibido / received: 27/05/2020. Aceptado / accepted: 16/06/2020.

1 Autor principal. Ingeniero en Diseño Industrial y Desarrollo del Producto.

2 Coautor. Doctor ingeniero industrial.

Autor para correspondencia: Asier Murcia Gila. E-mail: a.murcia.gila@gmail.com.



Introducción

El *big data* o análisis masivo de datos se ha convertido en uno de los fenómenos más populares en el avance tecnológico actual. Es fácil asociar este fenómeno con áreas informáticas, financieras y económicas. Sin embargo, pocas veces se visualiza como la gran herramienta de diseño que es.

Las nuevas tecnologías han cambiado la forma de actuar y trabajar, ofreciendo nuevas formas de entender el mercado, la demanda y la búsqueda de soluciones. El *big data* es una herramienta vanguardista para el análisis del entorno que rodea al diseño, una herramienta que permite analizar información en grandes cantidades y transformarla en modelos que ayuden en la toma de decisiones. Por ello, el análisis masivo de datos, así como sus nuevas formas de procesamiento de datos como el *machine learning* deben ser percibidas como una fase más en el desarrollo de un proyecto, una herramienta para comprender qué ocurre en un mundo cada vez más globalizado donde la información es un nuevo activo de gran valor.

'Big data' como herramienta de diseño

Un ejemplo práctico del potencial real del *big data* en el diseño industrial y

análisis del entorno lo encontramos en el sector del automóvil, con una transición de vehículos tradicionales de combustión a alternativas menos contaminantes como los vehículos eléctricos. La gran incertidumbre y el interés industrial que aborda el tema hacen que un incorrecto análisis de la situación actual pueda originar la toma de medidas precipitadas que causen una futura crisis energética o económica.

El estudio que se propone a continuación se ha realizado siguiendo una estructura general del *big data* construida con el objetivo de diversificar las diferentes fases que lo conforman, por lo que permite ser utilizada en múltiples aplicaciones.

Como todo análisis estadístico, el *big data* encuentra su primer escalón en el elemento más básico que lo forma, el dato. Las llamadas fuentes de datos, ya sean propias, de otras empresas o de fuentes oficiales, sustentan la fiabilidad de la mayor parte del proyecto. En *big data*, el gran volumen de los conjuntos de datos con los que se cuenta hace inviable la comprobación individual de cada uno de ellos. Por tanto, la fiabilidad de las fuentes de datos desempeña un papel importante en el proceso.

El resultado de todo ello se ve reflejado en caos. Y es que no hay otra forma

de definir la suma de toda la información recopilada de las fuentes de datos. El caos es algo inútil por sí mismo. Es necesario aplicar una gobernanza que permita limpiar y organizar los conjuntos de datos en el denominado *data lake*, un lago de datos ordenado, con formatos compatibles entre sí, del cual poder nutrirse. Finalmente, el procesamiento de los conjuntos de datos ya ordenados, mediante nuevas técnicas como el *machine learning*, permite obtener resultados de alto valor para el análisis del entorno y la toma de decisiones.

En el caso del estudio, para nutrir de suficiente información el *data lake*, se obtienen datos del sector del automóvil a través de la Dirección General de Tráfico (DGT), la Asociación Española de Fabricantes de Automóviles y Camiones (ANFAC), el Instituto Nacional de Estadística de España (INE) y de la Red Eléctrica Española (REE), esta última a través de su API oficial. Todos los conjuntos de datos se limpian y se transforman a formatos compatibles para permitir su procesamiento y se almacenan en una base de datos tipo NoSQL a través de Mongo DB.

El medio utilizado a lo largo del proyecto ha sido el lenguaje de programación Python, gracias a su simplici-

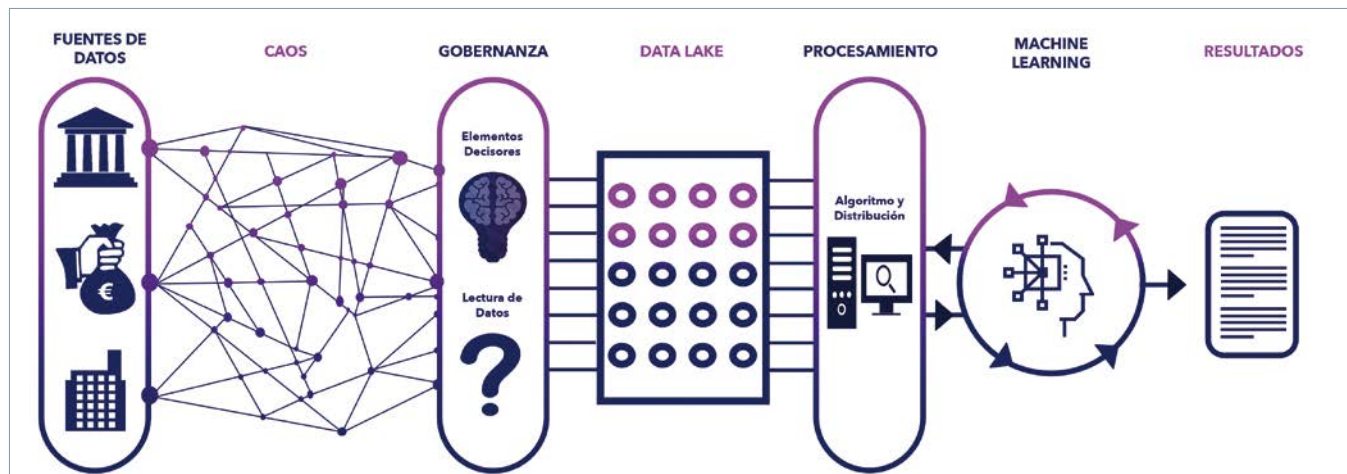
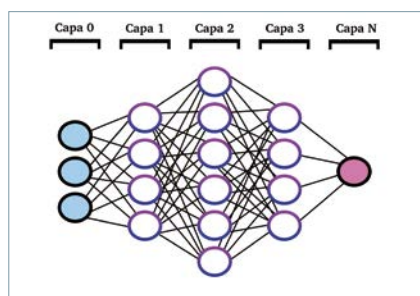
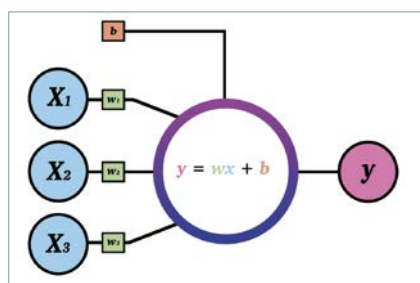


Figura 1. Estructura general del big data. Elaboración propia.



Figuras 2 y 3. Representación de una neurona y de una red neuronal. Elaboración propia.

dad y a la gran cantidad de bibliotecas con las que cuenta. Para realizar las predicciones planteadas en este proyecto, se hace uso de algoritmos como el método de suavización exponencial triple para estimar las tendencias de fenómenos en series temporales, a través de la biblioteca Fbprophet, y el uso de redes neuronales a través de la biblioteca Keras.

Redes neuronales para el procesamiento de datos

Las redes neuronales son el algoritmo más característico cuando se habla de *machine learning*. Diseñadas específicamente para esta función, permiten resolver problemas tanto de regresión como de clasificación. A diferencia del

resto de algoritmos tradicionales usados en estadística, las redes neuronales requieren de mucha más capacidad de procesado, y no ha sido hasta el siglo XXI cuando se ha contado con medios suficientes como para poder aplicar redes neuronales de forma masiva y al alcance de cualquiera.

La neurona tiene como objetivo codificar una información de entrada y obtener una información de salida, fenómeno similar a una función matemática en la que la introducción de unos valores X, ponderados con los valores W y un término independiente tiene como resultado el valor Y. La creación de múltiples neuronas organizadas en capas interconectadas permite generar sistemas neuronales capaces de procesar información con múltiples condiciones de estudio posibles. Para ello, se escalan los datos para la posterior aplicación de funciones de activación encargadas de separar los datos según las condiciones introducidas. La suma de todas ellas ofrece el resultado final del sistema neuronal. Para comprobar el error, el sistema neuronal hace uso del descenso del gradiente, obteniendo un vector que contiene las pendientes de cada una de las dimensiones de la función y se desplaza en sentido contrario.

Sin embargo, para poder hablar de aprendizaje automático, la red debe ser capaz de reajustar sus propios parámetros para corregir el resultado, y es ahí donde entra el algoritmo Backpropagation. Dada la inmensa cantidad de conexiones existentes en una red neuronal, la comprobación individual de cada una de ellas haría inviable el

proceso. Backpropagation realiza un estudio del error partiendo del resultado final y desplazándose capa por capa hasta encontrar las conexiones que generan el error. Esto permite descartar las conexiones que son correctas, y al ser un sistema interconectado, también permite descartar todas las anteriores a esta, lo que facilita la búsqueda del error.

En el caso del estudio propuesto se utiliza una red neuronal Long Short Term Memory (LSTM). Este tipo de redes recurrentes funcionan mediante aprendizaje supervisado, es decir, se debe suministrar a la red de datos reales para su entrenamiento. Son muy utilizadas en la predicción de fenómenos en series temporales, al recordar tendencias o dependencias de valores. A continuación se muestran los resultados del estudio.

Estudio de la transición a vehículos eléctricos en España

En la actualidad, hay numerosos informes acerca de la transición de vehículos de combustión a eléctricos. Debido a la gran complejidad del tema, es necesario delimitar correctamente los factores de interés para el objetivo de este estudio. Desde el punto de vista de interés industrial, los factores técnicos del vehículo, del mercado, así como los energéticos, cuentan con mayor peso.

El parque de turismo español ha crecido progresivamente hasta la actualidad, a excepción de la crisis económica de 2008 (figura 5). Sin embargo, si se atiende a la antigüedad del parque, según el último informe de Arval Mobility Observatory, el 62,1%

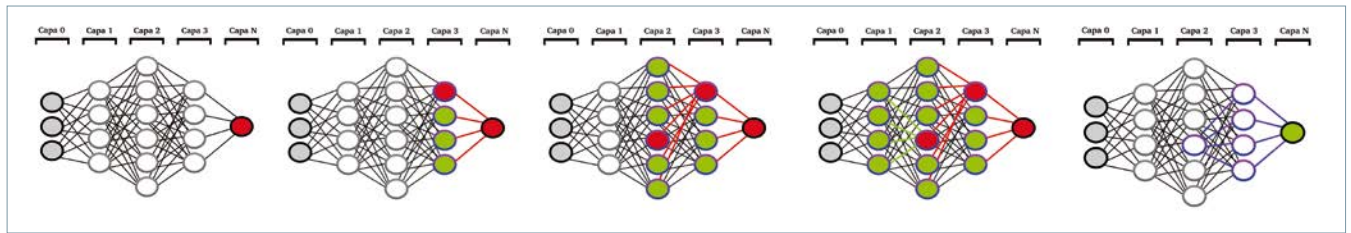


Figura 4. Algoritmo de Backpropagation. Elaboración propia.

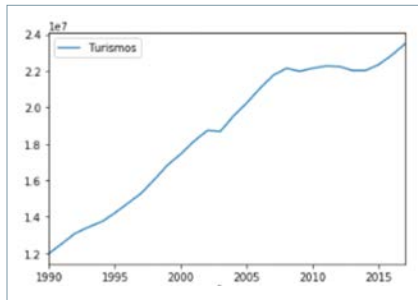


Figura 5. Serie histórica del parque de turismos español. Fuente: DGT.

de los turismos supera los 10 años de edad, dato superior a la media europea.

El dato gana valor cuando se observa que, de media, los vehículos españoles cuentan con una vida útil de 12 años de edad, es decir, en la próxima década más de la mitad del parque de turismos español deberá renovarse. A la hora de la renovación de turismos, el principal factor que hay que tener en cuenta es el precio. Según la información publicada por la Agencia Tributaria de España (figura 6), se observa que las ventas de vehículos pertenecientes a gama baja (<15.000 €), así como los de gama media (15.000-30.000 €), son muy superiores a los de gama

alta (>30.000 €). Según un artículo publicado por *Autonomi* en 2018, el precio medio de los vehículos eléctricos a la venta en ese año se situaba en 49.986 €, muy alejado de las franjas de precio más vendidas.

Atendiendo a la división del parque de turismos español por combustible, los vehículos diésel, junto con los de gasolina, abarcan el 95% del total. El 5% restante, denominado vehículos ECO, incluye a los híbridos no enchufables (HEV), híbridos enchufables (PHEV), propulsados a gas (GLP) y eléctricos de batería (BEV). Estos últimos no suman ni el 1% del total (figuras 7 y 8).

El crecimiento de este tipo de vehículos ECO, menos dañinos con el medio ambiente, viene impulsado desde las medidas adoptadas en el año 2016 por parte del Gobierno que supusieron el mayor paquete de medidas contra vehículos contaminantes, con la implantación del distintivo medioambiental de la DGT. Ejemplo de ello es la caída en ventas de los vehículos diésel después de su prohibición en los centros de las principales ciudades del territorio español. De igual forma, también supuso un punto de inflexión

en el crecimiento en ventas de vehículos de gasolina y ECO.

Ese aumento de los vehículos de gasolina encuentra explicación en una encuesta realizada por la plataforma coches.net, en la que se puso en manifiesto la penalización en el mercado del alto precio de los vehículos híbridos y eléctricos. De hecho, el 35% de los encuestados compraría un BEV en caso de que el precio de estos bajase. El porcentaje restante demandaría también mejora en otros aspectos: el 31% una mejora de la autonomía, el 25% resalta la dificultad para encontrar puntos de recarga, el 11% gastos de mantenimiento de las baterías, el 10% la poca oferta existente y el otro 10% el alto precio que supone el cambio de las baterías.

En conclusión, se puede definir que la reducción del precio de los vehículos eléctricos, el desarrollo de su tecnología y la mejora de su autonomía y vida útil de las baterías serán clave para su crecimiento. Igualmente, las medidas que adopten los gobiernos tendrán importantes impactos en las tendencias de ventas de cada modelo.

Tras haber realizado un análisis general del entorno que rodea la transición a vehículos eléctricos, se pueden construir tres escenarios diferentes, permitiendo así poner en contexto los resultados obtenidos:

- **Escenario continuista:** asume que no va a producirse ningún cambio notable en la legislación, manteniendo la tendencia originada en 2016 en cuanto a la compra de vehículos. De igual forma, supone una evolución tecnológica a ritmo normal y la llegada por inercia a la renovación del parque.
- **Escenario pesimista:** asume un retroceso en la evolución del parque, originada por múltiples factores que pueden ir desde crisis económicas hasta estancamiento en el desarrollo de la tecnología, lo que retrasaría

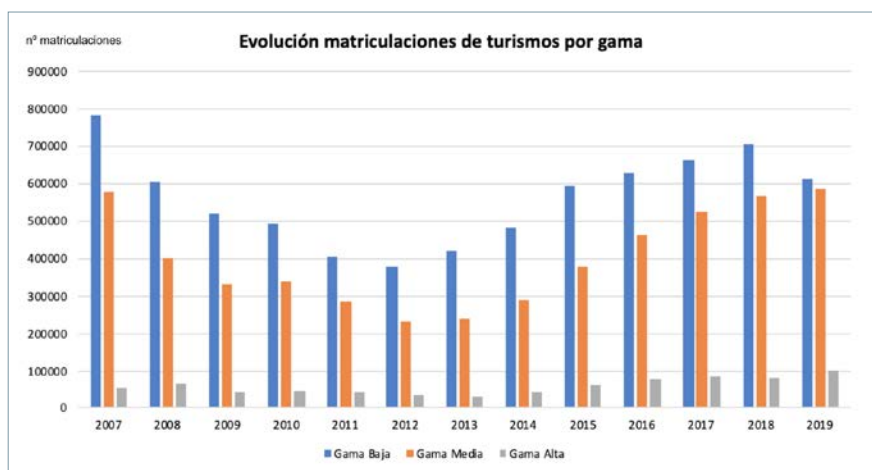


Figura 6. Evolución de las matriculaciones de turismos por gama. Fuente: Agencia Tributaria de España.

la renovación del parque.

- **Escenario optimista:** asume una evolución notable originada por la necesidad de renovación del parque de turismos español. De igual forma, se asumen futuras normativas que incentiven la venta de vehículos ECO y que penalicen a los más contaminantes. La evolución tecnológica de los BEV será clave para lograr reducir su precio y mejorar sus prestaciones, fomentando así su compra.

Bajo estos tres escenarios, se procesan los datos de tendencia de crecimiento del parque vehicular y de compra general de turismos, obtenidos de la DGT (figuras 8 y 9), así como la tendencia según combustible, obtenida del estudio realizado por Anfac, Faconauto y Ganvam. Se hace uso de la biblioteca Fbprophet en Python:

Los resultados del crecimiento del parque de turismos español mantienen la progresión de los últimos años. Sin embargo, se debe tener en cuenta la aparición de nuevos medios de transporte que sirven de alternativa al vehículo privado, sobre todo en las grandes ciudades, y que ralentizan el crecimiento. En cuanto a la predicción de ventas, tanto el escenario continuista como el más optimista prevén un crecimiento en los próximos años, fruto de la necesidad de renovación de gran parte del parque. Sin embargo, la posibilidad de una crisis económica que afecte al sector del automóvil también debe tenerse en cuenta, y se escenifica en la vertiente más pesimista de las tres, con una caída notable en ventas que ralentice el crecimiento del parque. A continuación, se muestran los resultados de tendencia de ventas por

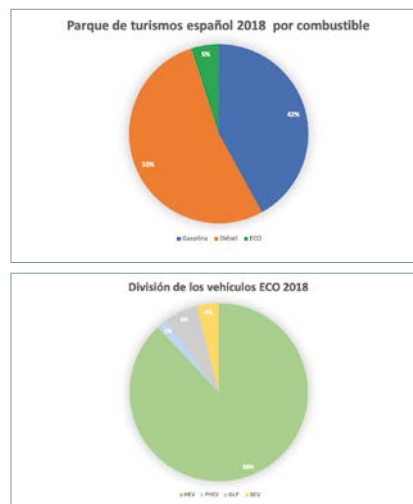
combustible (los porcentajes de venta de vehículos ECO incluyen también a los BEV, se diferencian para poder observar la incidencia individual de estos últimos sobre el total de ventas) (figuras 10-12).

El escenario continuista, fruto de las medidas ya adoptadas y las tendencias que estas generaron a partir del año 2016, colocan a los vehículos ECO como los grandes beneficiados, con un importante crecimiento en ventas a lo largo de los años. Sin embargo, este crecimiento viene de la mano de vehículos híbridos, los cuales, al ser una transición lógica entre vehículos de combustión y alternativas eléctricas, combinan beneficios de ambos. Los tres escenarios certifican la caída de los vehículos diésel, así como cambios mínimos para los vehículos gasolina, los cuales no pierden nicho de mercado.

En el escenario más optimista, al suponer medidas por parte de los gobiernos que influyan también sobre los vehículos gasolina, sitúa a los ECO al frente del mercado. En este caso, y a diferencia del escenario continuista, el desarrollo en la tecnología de los BEV, junto con una reducción de su precio, los convertiría en la alternativa más vendida.

Esta parte del estudio permite observar la evolución del mercado en función de los tres escenarios. A continuación, se muestra el impacto que tendrían estas tendencias de venta sobre el parque total de turismos (figuras 13-15).

Como se puede observar, los resultados reflejan una caída progresiva de los vehículos de gasóleo en los tres escenarios, aunque la renovación es lenta a causa del gran peso de estos vehícu-



Figuras 7 y 8. Distribución del parque de turismos español en 2018 por combustible. Fuente: DGT.

los en el parque actual (53%). Por su parte, los vehículos de gasolina logran mantenerse notablemente en los tres escenarios, como ya se ha visto en las tendencias de venta.

En el escenario continuista, el *sorpasso* de los vehículos de bajas emisiones a los vehículos diésel no llegaría hasta aproximadamente 2038. Por su parte, en el escenario optimista, los vehículos ECO lograrían convertirse en los dominadores generales del parque español, por encima incluso de los vehículos de gasolina. Es el caso contrario que en el escenario pesimista, originado por la baja cantidad de ventas, que limitaría el crecimiento de los ECO a tan solo el 18% del total en 2040, el 7% si se cuenta solo BEV.

El principal objetivo de la transición es la reducción de emisiones de CO₂ a la atmósfera, y para lograrlo,

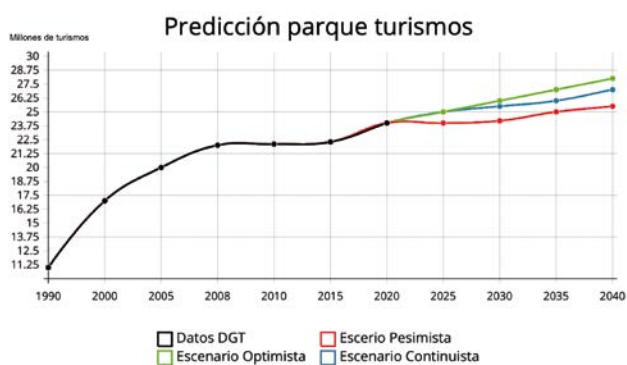


Figura 8. Predicción de la evolución del parque de turismos. Elaboración propia. Datos de entrenamiento: DGT.

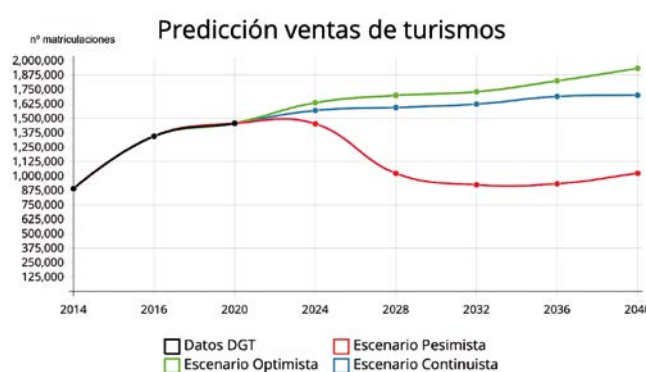
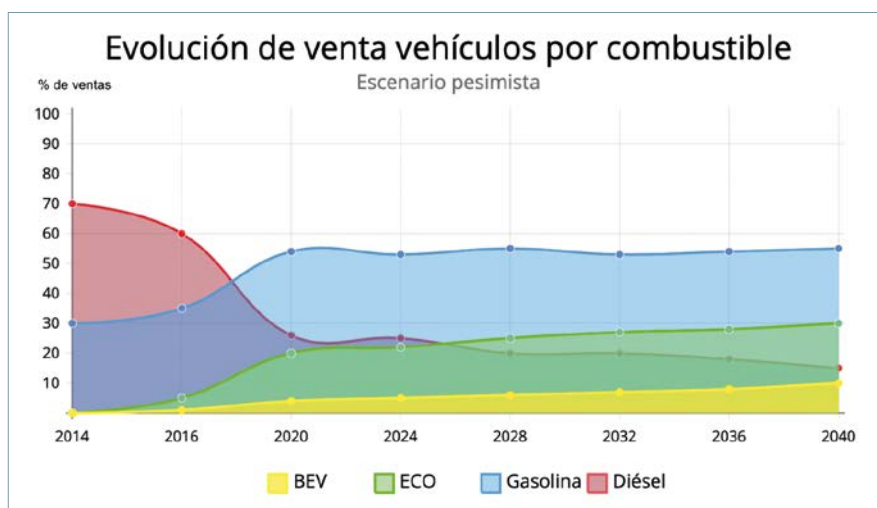
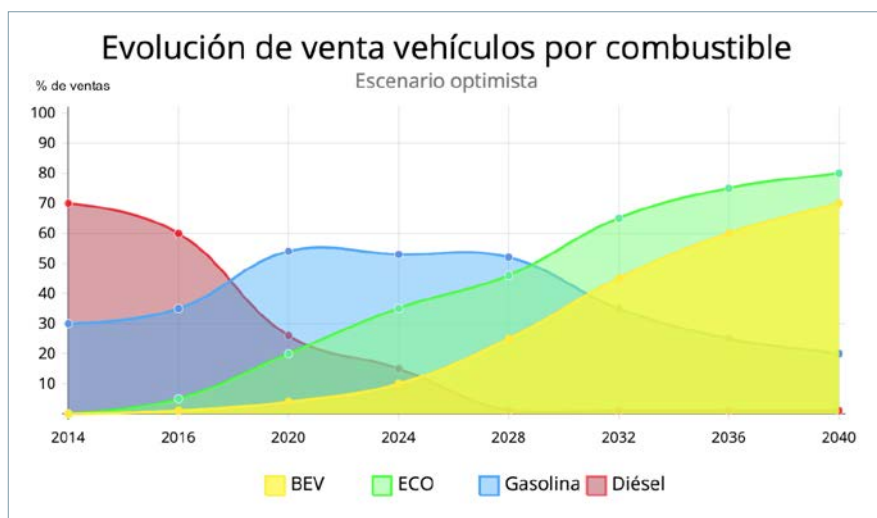
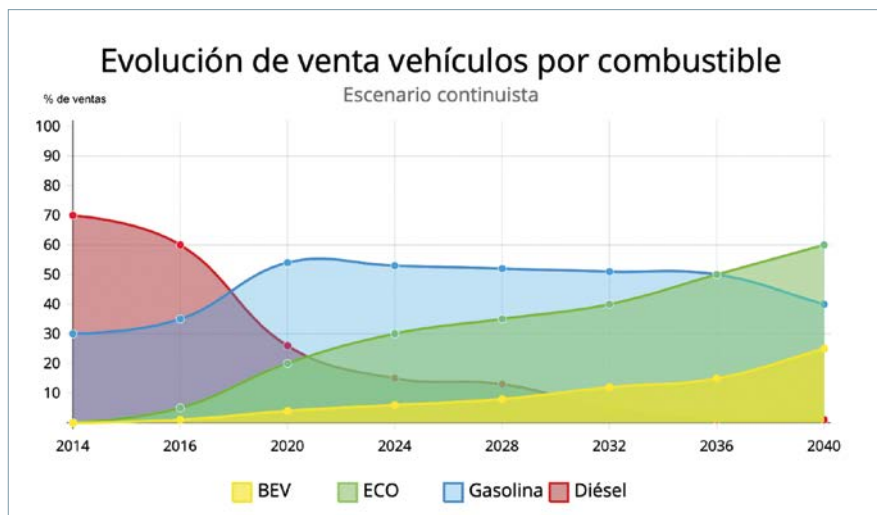


Figura 9. Predicción de la evolución de las ventas de turismos. Elaboración propia. Datos de entrenamiento: DGT.



Figuras 10-12. Evolución de la venta de vehículos por combustible. Elaboración propia. Datos de entrenamiento: ANFAC, Faconauto y Ganvam.

será necesario que la energía eléctrica requerida para cargar los BEV provenga de fuentes limpias. En España, a fecha de 2020, se cuenta con una capacidad de generación máxima de 110.000 MW, de los cuales 55.000 MW son

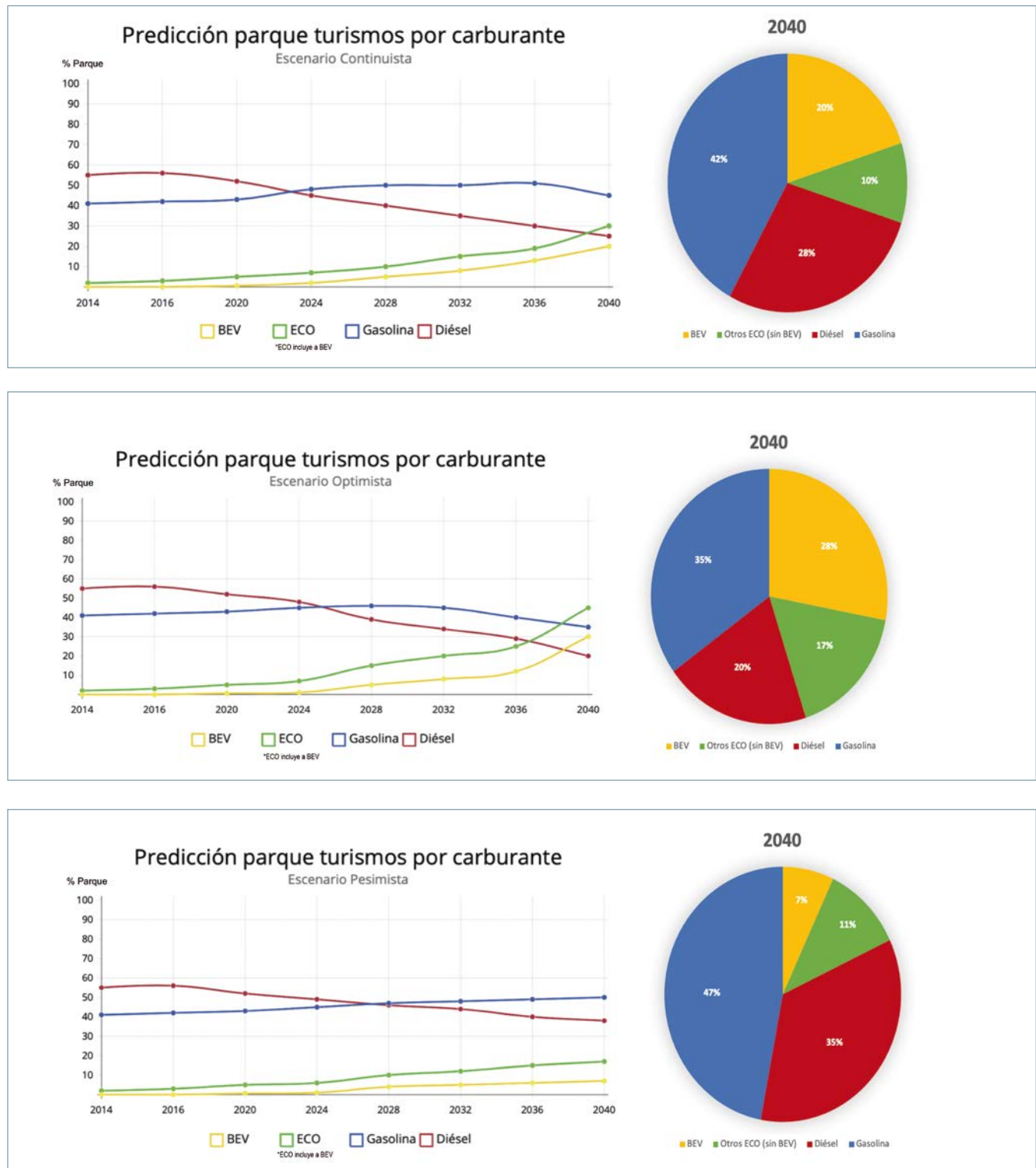
de energía renovable. Para estudiar el impacto que tendrían los diferentes escenarios planteados en la red eléctrica, se construye una red neuronal LSTM haciendo uso de los valores de demanda horaria desde 2010 hasta 2019 para

su entrenamiento, obtenidos a través de la API de Red Eléctrica Española (figura 16).

En esa predicción, el tipo de carga del vehículo, así como el índice de simultaneidad de las conexiones, cuentan con un gran peso. Para este estudio, se opta por una potencia media de 12 kW por carga, atendiendo a los tipos de cargadores actuales, y a la evolución de los mismos. Por su parte, teniendo en cuenta la autonomía media de los BEV actuales (200 km de media reales, según *Autonomión*) y la media de kilómetros recorridos por día (35 km, pese a que en las grandes ciudades es mayor) se opta por una frecuencia de carga cada 3 días. La conexión de los vehículos a la red se realizaría de forma progresiva.

Se seleccionan los resultados del escenario optimista anterior, al contar estos con el mayor número de vehículos eléctricos (7 millones en 2040). Se plantean tres nuevas formas de efectuar la carga: una concentrada en las horas valle de menor demanda, una dividida en horas valle y horas de jornada laboral (al no contar toda la población con posibilidad de cargar los vehículos en sus hogares) y una tercera carga ineficiente coincidiendo con las horas pico, en la vuelta al hogar tras la jornada laboral. En los tres casos, se escenifica que ocurriría con una carga de 1/3 del total de BEV, 2/3 y 3/3. Los resultados de la red neuronal fueron los siguientes (figuras 17-19):

En conclusión, se puede observar que, con la capacidad de generación actual, a fecha de 2020, y suponiendo una flota 7 millones de BEV, no se pone en peligro el sistema eléctrico español ni en el peor de los escenarios, causa directa de la baja demanda media existente en el país. Sin embargo, el problema aparece al observar cuánto porcentaje de esa energía proviene de fuentes limpias. En 2020, sería inviable hacer frente a tal cantidad de BEV solo con energías renovables. Si se observan los planes del Gobierno español, en la propuesta de ley para el cambio climático y la transición ecológica, se establece un plan de instalación de 3.000 MW de potencia renovable de forma anual, lo que equivaldría a 60.000 MW nuevos en 2040, unos 115.000 MW en total. En este escenario, hasta en la carga más ineficiente se contaría con capacidad renovable. A este análisis



Figuras 13, 14 y 15. Predicción del parque de turismos por carburante en tres escenarios. Elaboración propia. Datos de entrenamiento: ANFAC, Faconauto y Ganvam.

sis hay que añadir el avance tecnológico de las baterías, puntos de recarga y fuentes de generación que optimizarán el sistema, reduciendo la cantidad de potencia en cada carga, así como la demanda general, con lo que se puede llegar a hacer frente a una cantidad de BEV muy superior. Sin embargo, hay

muy poco margen de error, por lo que la apuesta por la construcción de plantas de generación energética renovable debe ser firme, si se desea una transición exitosa.

Desde un punto de vista de interés industrial, y atendiendo a los resultados de tendencias de venta, los vehícu-

los híbridos cuentan con la mejor proyección a corto plazo, al igual que los vehículos gasolina, estos últimos, dado su dominio en franjas de bajo precio, al menos hasta 2040. Por su parte, los vehículos diésel dejarán de contar con un nicho de mercado rentable sobre 2030. De igual forma, el éxito de los

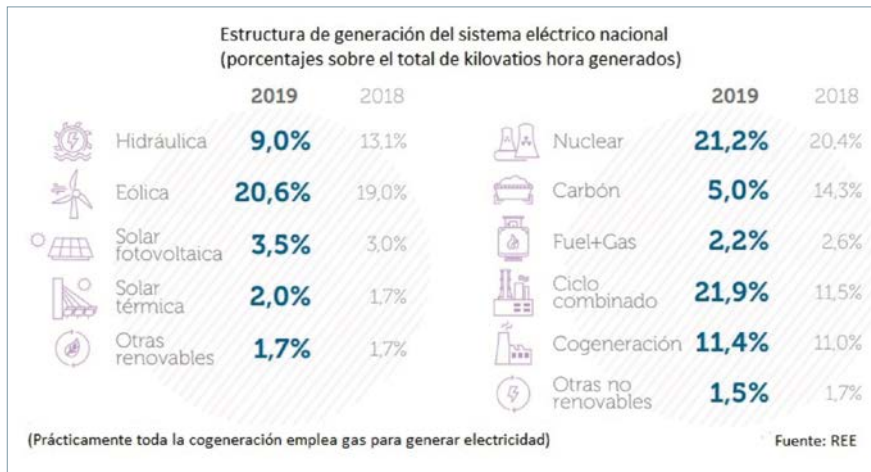
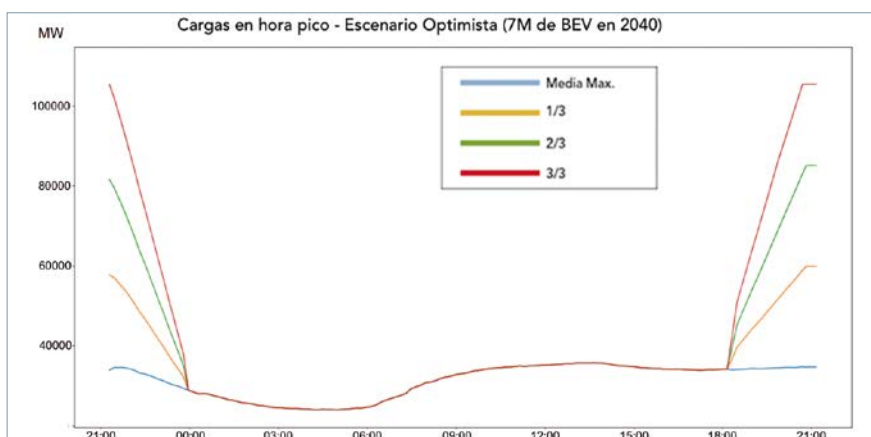
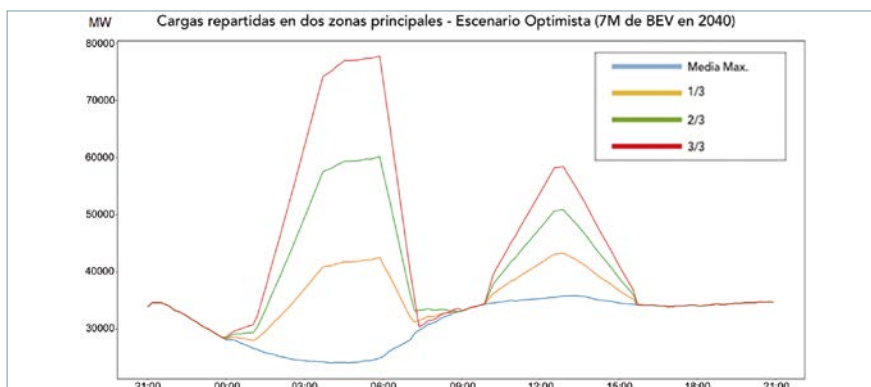
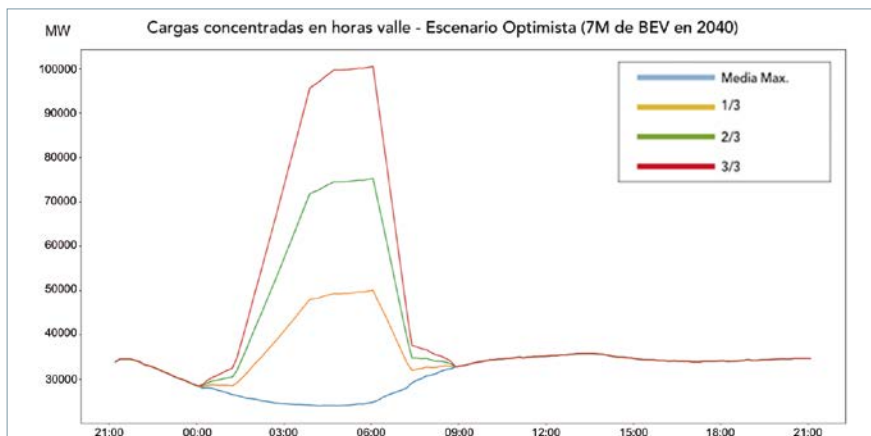


Figura 16. Estructura de generación del sistema eléctrico español. Fuente: REE.



Figuras 17, 18 y 19. Impacto carga BEV en la demanda horaria. Elaboración propia. Datos de entrenamiento: REE.

vehículos eléctricos dependerá principalmente de su desarrollo tecnológico y su reducción del precio de venta, permitiendo ser competitivos en nichos de 10.000-20.000 €.

Las conclusiones del estudio no se limitan únicamente al sector del automóvil. Como se ha podido observar, la gestión de la demanda eléctrica también afectará al desarrollo de la transición. Por ello, la reducción de potencia demandada a la red por parte de otros productos como electrodomésticos comenzará a tener gran importancia y valoración por parte del cliente, ya no solo por el ahorro económico que supone, sino por el ahorro energético que permitirá cargar los BEV de forma más rápida y efectiva.

Este estudio forma parte del trabajo de fin de grado: *Diseño basado en 'big data' para aplicación en ingeniería*, presentado en la Escuela Técnica Superior de Ingeniería y Diseño Industrial de la Universidad Politécnica de Madrid. En él, se exponen de forma extensa y didáctica todas las fases del *big data* para su aplicación en diseño industrial, ayudando a comprender el entorno que lo rodea, conjunto al estudio práctico expuesto en este artículo.

La era del *big data* no ha hecho más que comenzar, y es importante comprender que esto es solo la cúspide de todo un iceberg que se irá desarrollando en los próximos años, que cambiará por completo la forma de comprender el entorno y que es la piedra angular de la ingeniería del futuro.

Referencias

- Informe del parque vehicular español (2019). Arval Mobility Observatory. Disponible en: https://www.arval.es/sites/es/files/media/pdfs/arval_amo_2019_interactivo.pdf
- Matriculaciones por tramo de venta (2018). Agencia Tributaria de España.
- Vehículos eléctricos a la venta (2018). Autonoción. Disponible en: <https://www.autonocion.com/coches-electricos-precio-espana/>
- Parque de turismos (2018). Dirección General de Tráfico.
- Evolución de las ventas turismos por combustible (2020) ANFAC, Faconauto y Ganvam. Disponible en: <https://www.epdata.es/tjyujt/b68e3d97-f49e-46d0-87b8-594ecf46d4e8/espana/106>
- Sistema de generación de energía (2019). Red Eléctrica de España.
- Kilómetros diarios recorridos por un turismo (2018). Instituto Nacional de Estadística.
- Ley de cambio climático y transición ecológica (2020). Ministerio para la Transición Ecológica.