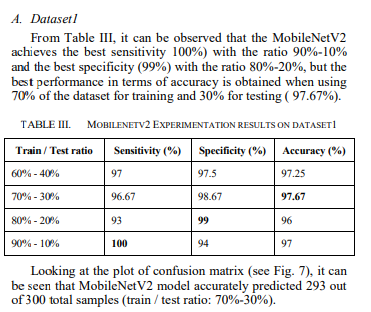
# TRAIN TEST split comprobado

<https://www-proquest-com.universidadviu.idm.oclc.org/docview/2992551140?pq-origsite=summon&sourcetype=Scholarly%20Journals>



# 1.- Urban Residential Energy Demand and Rebound Effect in China: A Stochastic Energy Demand Frontier Approach

<https://www.proquest.com/docview/2542356333?parentSessionId=EVh9uH%2FeCJIqa6%2Bb8EBMRkKj51l79uQfBrAxKWkwcwE%3D&pq-origsite=summon&accountid=198016&sourcetype=Scholarly%20Journals>

se puede citar algo de los factores que influyen en la demanda de la energía

2.- A PSO–GA optimal model to estimate primary energy demand of China

<https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0301421511009888>

también hay factores que influyen en la demanda.

Energy demand modeling is an essential component for energy planning, formulating strategies, and recommending energy policies. Forecasting energy consumption is generally difficult because it is affected by the rapid development of the economy, technology, government decisions, and other factors. Thus, the accurate estimation of energy demand is a very critical factor in domestic energy policy making.

### 2.2. Artificial intelligence models

The second group is in the field of artificial intelligence, including artificial neural networks (ANN), ant colony optimization (ACO), genetic algorithms, and particle swarm optimization (PSO) algorithm. The details are as follows:

* (1)  
  *Artificial Neural Networks forecasting model.* The ANN technique may be considered a regression technique that represents higher nonlinearity between independent and dependent variables ([Geem et al., 2007](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0301421511009888#bib26)). ANN models are inspired by the biological neural system, with capability to learn, store, and recall information based on a given training dataset. ANN models are “black-box” modeling techniques, which are capable of performing non-linear mapping of a multidimensional input space onto another multidimensional output space without the knowledge of the dynamics of the relationship between input and output spaces.

## 4. Factors affecting demand and data management

### 4.1. Factors affecting demand of China

3.- Modeling Energy Demand—A Systematic Literature Review

s. A medida que el alcance de estos modelos se amplía a través de múltiples infraestructuras y portadores de energía [1], se vuelven cada vez más detallados y complejos [2]. Por lo tanto, disponer de información bien fundamentada sobre la demanda futura de energía, con una alta resolución temporal y espacial, es uno de los insumos más importantes para estos modelos, ya que tiene un impacto directo en los procesos de toma de decisiones asociados [3], que afectan al funcionamiento de la red en tiempo real, así como a la planificación de la ampliación de la infraestructura a largo plazo. En 2020, este número había aumentado a 641. Los modelos de demanda de energía tienen una amplia gama de aplicaciones. Como muestran Bhattacharyya y Timilsina [6], pueden abarcar desde la previsión del consumo de energía a corto plazo en las redes y los mercados energéticos hasta la simulación de las cargas de calor y electricidad en edificios y procesos industriales, hasta proyecciones econométricas a largo plazo de la demanda energética nacional.

En [4], Debnath y Mourshed presentan una revisión sobre las técnicas de previsión de la oferta y la demanda en los modelos de planificación energética en todos los operadores de energía. Los autores presentan 483 modelos de artículos publicados entre 1985 y 2017. Analizan la extensión geográfica, los plazos y las medidas de rendimiento, así como los criterios específicos de las técnicas, como el número de neuronas en las capas de las redes neuronales artificiales (ANN). Si bien esta revisión proporciona un análisis amplio, no se incluyen los aspectos relacionados con los datos y no se hace una distinción entre los sectores.  
 Wei y otros [14] compilaron un estudio bibliográfico sobre modelos convencionales y basados en inteligencia artificial para la predicción del consumo de energía. Se han descrito 116 publicaciones con respecto al propósito, los horizontes temporales,

Hong y Fang [5] sugieren las dos categorías generales de técnicas estadísticas y de inteligencia artificial: la primera comprende las técnicas de regresión lineal múltiple y la TSA, y la segunda incluye la ANN, la regresión difusa, las máquinas vectoriales de soporte (SVM) y las máquinas de aumento de gradientes. Wei y otros [14] distinguen entre las técnicas convencionales, como la TSA, la regresión y los modelos grises, y las técnicas de inteligencia artificial, como la ANN y la SVM. Kuster y otros [9] analizan las categorías de la TSA, los modelos de regresión, la ANN, la SVM y las técnicas de abajo hacia arriba.

**3.2. Técnicas de aprendizaje automático**

Las técnicas de esta categoría encuentran una amplia aplicación en el modelado y la predicción de la demanda de energía y se pueden dividir en enfoques de aprendizaje supervisado y no supervisado.

Los enfoques de aprendizaje supervisado utilizan conjuntos de datos de entrenamiento etiquetados para derivar una función que describe una relación entre entradas y salidas basándose en ejemplos de pares de entradas y salidas [31]. Se pueden aplicar a variables numéricas en el caso de problemas de regresión y a variables categóricas en el caso de problemas de clasificación [32]. Dentro de esta subcategoría, las técnicas más comunes son la ANN y los algoritmos basados en instancias, como las máquinas k-near neighbor y kernel [33]. Un ejemplo común de estas últimas es la SVM, que puede convertir problemas no lineales en un espacio de baja dimensión en problemas lineales en un espacio de alta dimensión [34]. Además, en esta categoría hay árboles de decisión, algoritmos bayesianos y enfoques de aprendizaje por conjuntos, como las máquinas que aumentan el gradiente

[5].

Escuchar selección

Los enfoques de aprendizaje no supervisados a menudo se aplican a los problemas de agrupamiento. Estos algoritmos deducen estructuras en un conjunto de datos de entrada sin etiquetar, por ejemplo, mediante la búsqueda de similitudes

El país con la mayor producción de artículos durante los últimos años es China (98), seguido de EE. UU. (40) y Turquía (31). El análisis de las fechas de publicación muestra un ligero aumento de artículos en los últimos seis años, pasando de 66 en 2015 a 77 en 2020.

Según Hong y Fan, la medida de rendimiento más utilizada en la industria de la energía eléctrica es el MAPE, debido a su simplicidad y transparencia [5]. El índice de referencia de Lewis [127], mencionado por varios autores [18.128], sugiere que un valor de MAPE del 10% o inferior indica una alta precisión de

predicción.

En varias revisiones de la literatura se comparan los valores de MAPE de diferentes técnicas [4,14,23]. Debnath y Mourshed sugieren que el aprendizaje automático y los enfoques híbridos tienden a ser más precisos en comparación con otras técnicas [4], y Wei et al. descubrieron que los valores del MAPE de las proyecciones a largo plazo tienden a ser mejores que los de las proyecciones a corto plazo [14]. Sin embargo, otros autores se muestran reacios a dar recomendaciones claras, afirmando que las diferentes opciones de medidas de rendimiento dificultan la categorización de los métodos de mejor a peor [4] y que la idoneidad de los modelos depende finalmente del conjunto de datos [23  
Para las técnicas de aprendizaje automático, en los artículos analizados se han encontrado las siguientes medidas para mejorar el rendimiento predictivo. Para reducir el sobreajuste, se empleó el aprendizaje conjunto para crear predicciones independientes de múltiples modelos y utilizar resultados promediados ponderados [59,95,97,99,113,115,129,130,131,132,133,134]. Otras medidas contra el sobreajuste incluyen el uso de redes neuronales dinámicas y de aprendizaje incremental, donde los modelos se actualizan paso a paso durante la fase de entrenamiento [88,106,131,135] o se implementan restricciones en los coeficientes [136], así como la introducción de capas de abandono [137]. Los ajustes de los coeficientes de las variables de predicción para capturar las propiedades esenciales de los datos de entrenamiento y proporcionar una mejor generalización a puntos de datos aún desconocidos son un concepto importante y generalizado para evitar el sobreajuste conocido como regularización. La idea es utilizar una técnica de regresión para reducir o regularizar los coeficientes estimados, lo que desalentaría de manera efectiva el aprendizaje de un modelo complejo y, por lo tanto, reduciría el riesgo de sobreajuste. Los procedimientos más comunes son el operador de selección y contracción mínima absoluta (LASSO) [138,139,140,141,142] y la regresión de crestas [96,102,143,144]. La regularización bayesiana

se utilizó en [145,146].

La maldición de la dimensionalidad [147] es un problema común que se produce cuando se trata de conjuntos de datos de alta dimensión. La eliminación de las características insignificantes y la selección adecuada son cruciales para las técnicas de regresión y aprendizaje automático. Esto se logró mediante el análisis de correlación y de componentes principales (PCA) en [61,107,148,149,150,151,152,153,154,155,156,157,158,159,160,161]. En el caso de los enfoques de aprendizaje automático, la selección de características también se llevó a cabo mediante algoritmos genéticos y árboles de decisión en [39,40,100,111] y una selección de características de base difusa en [162,163,164]. Para aumentar la precisión, se utilizan técnicas de aprendizaje profundo, como los codificadores automáticos apilados [21.165.166.167] o las redes de memoria a corto plazo (LSTM) [153.168.169.170.171]. Otra medida para aumentar la precisión de la ANN es variar el número de neuronas y capas, como se muestra en [106,159,172,173

].Otra medida popular es el procesamiento previo de los datos para eliminar los valores atípicos y el ruido, así como para aislar los patrones estacionales [113] o relacionados con la temperatura [174]. En el caso de las técnicas TSA y ML, esto se ha realizado mediante la transformación de ondículas o de Fourier en [39,58,90,108,111,120,175,176,177,178,179,180]. Se prestó especial atención a la predicción de eventos especiales y días festivos en [

118,181,182,183].

A veces, el rendimiento del modelo se ve afectado por la falta de datos. En el caso de los enfoques de aprendizaje automático, esto se puede compensar con la creación de datos virtuales mediante funciones de densificación o de información latente [184,185]. En las técnicas basadas en la ingeniería, la insuficiencia de datos puede abordarse mediante la priorización y la elección correcta de muestras representativas, como se hizo en [19,54,186,187,188

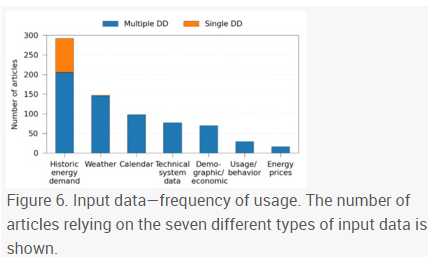
].

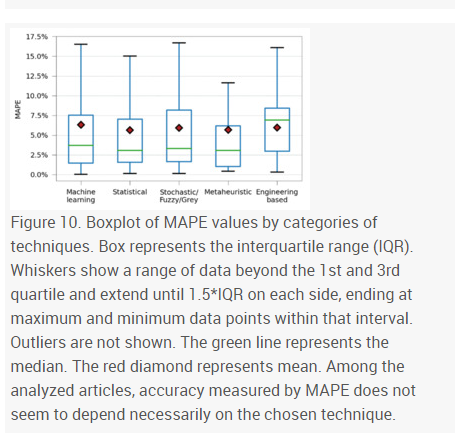
Los enfoques de ML a veces no terminan en mínimos locales poco profundos al optimizar los parámetros. Las soluciones pueden incluir rutinas de optimización alternativas (metaheurísticas) durante la fase de entrenamiento, como el algoritmo de colonias de abejas artificiales para las ANN [116,189], la búsqueda Cuckoo [71] o el algoritmo de manada de lobos [163]

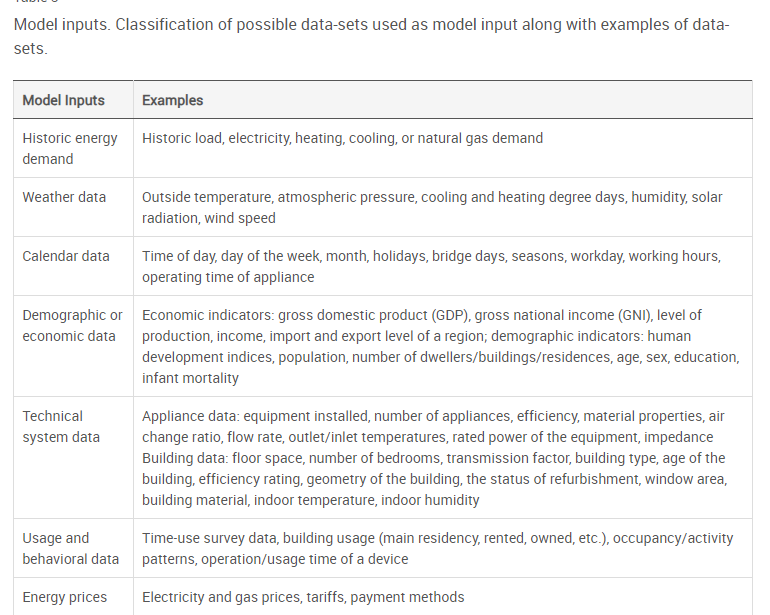
Las técnicas de aprendizaje automático son las que más se utilizan en los artículos y mostraron un aumento en el uso en los últimos años. El aprendizaje automático tiene la ventaja de poder manejar relaciones no lineales y lograr altos niveles de precisión con un esfuerzo de implementación bastante bajo [22]. Los inconvenientes radican en el carácter de caja negra [190], en la tendencia a sobreajustarse y quedarse atascados en mínimos locales poco profundos [116.191]. Las contramedidas son el uso de procedimientos de regularización, la formación de conjuntos de modelos, la selección de características y el preprocesamiento de datos por descomposición. Las técnicas de aprendizaje automático predominan en todos los niveles temporales y espaciales, aunque con una ligera tendencia a reducir los intervalos de tiempo, los horizontes y las escalas  
Teniendo en cuenta los intensos esfuerzos realizados en relación con los objetivos de eficiencia y el potencial de gestión de la demanda, existe un gran interés en modelar el sector industrial, incluida la adopción de nuevas tecnologías, como explican Fleiter y otros [76]. El reducido número de artículos puede explicarse por la falta de datos disponibles públicamente, así como por la renuencia a publicar este tipo de investigaciones; el consumo de energía de una empresa a menudo se considera un dato confidencial, ya que contiene implícitamente información sobre la actividad de producción y la eficiencia [125]. Los resultados de la revisión en cuestión indican que persisten los desafíos señalados por Fleiter y otros en relación con la disponibilidad de datos y la transparencia en el sector industrial [76].

La

precisión de la predicción es uno de los factores más importantes en la toma de decisiones, no solo para permitir la elección correcta de los modelos, sino también para permitir que las partes interesadas comprendan el rendimiento del método empleado. A diferencia de los hallazgos de otros autores ([4,14]), en los artículos analizados no se puede confirmar la tendencia a una mayor precisión con las técnicas híbridas y de aprendizaje automático o con horizontes temporales más largos.







5. Forecasting China’s regional energy demand by 2030: A Bayesian approach

<https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0921344917302550>

China has been the largest energy consumer in the world, and its future energy demand is of concern to police makers due to the significance for strategic planning. In 2015, China’s energy consumption totaled 4.30 billion tons of standard coal equivalent (SCE) of which coal accounted for 64.0%.

<https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0301421515302329>

In this study, both approaches are combined in a way that, the future values of the descriptor variables (i.e. population) were predicted using the past values of these variables by time series ANN models, and the future values of the electricity demand were forecasted using the predicted values of the descriptor variables by multilayer perceptron ANN models. Inspired from the biological nervous systems, ANNs mimic the learning that happens in humans; they have a great ability to approximate any nonlinear relationship that exists between a set of input variables and an output variable

Artificial neural network (ANN) models were used for two different purposes; for the ANN-Type I (multilayer perceptron), the electricity demand was modeled as a function of the descriptor variables as shown in Fig. 2 (only the statistically significant variables were used as input variables in the final model). Next, in order to simulate this ANN model for the future years, the future values of the statistically significant descriptor variables (i.e. population) was determined by using time series ANN models (ANNType II), in which one particular variable in the year “t” was modeled as a function of its values in the past years. The conceptual ANN model of this type is shown in Fig. 3 (for the population as an example).

Two hidden layers were used in the ANN architectures to capture any possible nonlinearity in the data more effectively

6.- Deep neural network based demand side short term load forecasting

<https://www.proquest.com/docview/1862118259?parentSessionId=eBegW49o5zEKVh27sxHOqOd0pAe2EubYz91S%2BVnLPrA%3D&pq-origsite=summon&accountid=198016&sourcetype=Scholarly%20Journals>

7.- Building Energy Time Series Data Mining for Behavior Analytics and Forecasting Energy consumption

<https://web-p-ebscohost-com.universidadviu.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=0&sid=6126dbd5-f241-4384-958a-55a0d487e027%40redis>

8.- Long-Term Demand Forecasting in a Scenario of Energy Transition

<https://www.proquest.com/docview/2317018068?parentSessionId=Ih7DfkjJoHZxQDzaJNKNIuiZDE1y8afr59rKuSy8HZA%3D&pq-origsite=summon&accountid=198016&sourcetype=Scholarly%20Journals>

sirve más para introduccion

9.- A prediction model based on neural networks for the energy consumption of a bioclimatic building

<https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349>

Artificial intelligence methods are a research line that has been experiencing an increasing focus over the past years because of their good fit to this kind of problems. Artificial intelligence includes several techniques as [artificial neural networks](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/artificial-neural-network), fuzzy logic, [support vector machine](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/chemical-engineering/support-vector-machine), genetic programming or a combination of them which is also known as hybrid system. Neural networks are widely used for the task of building energy consumption prediction because they are good at solving non-linear problems and fit well to this complex problem [[21]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0105), [[22]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0110), [[23]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0115). Fuzzy logic is a form of probabilistic logic that deals with reasoning that is approximate rather than fixed or exact, fuzzy logic's success in these applications has been attributed to its ability to effectively model real world data [[24]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0120), [[25]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0125). Support vector machines are increasing their presence in research and industry over the last years, they are effective dealing with non-linear problems even with small historical data sets [[26]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0130), [[27]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0135), [[28]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0140).

. The choice of neural networks for this task has been made because of the complexity of the analyzed problem and the ANNs distinctive features: learning, self-adaptative, fault tolerance, flexibility and real [time response](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/response-time) [[34]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0170), [[35]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0175), [[36]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0180).

* 1.  
  Data collection2.  
  Variable preselection
* 3.  
  [Data preprocessing](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/data-preprocessing)
* 4.  
  Training, validation and testing sets construction
* 5.  
  Neural networks paradigms
  + (a)  
    Architecture
  + (b)  
    Structure
* 6.  
  Variable selection
* 7.  
  Model order determination
* 8.  
  [Neural network training](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/neural-network-training)
* 9.  
  Evaluation and implementation

### Neural network model

Artificial neural networks (ANNs) have represented a new paradigm in the energy prediction. They work as a black-box model, thus, it is not necessary to have detailed information about the system because ANNs learn the relationship between input and output parameters by means of [historical data](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/historical-data). Another advantage is the ability to manage large complex systems with so many parameters interrelated between themselves, since they may ignore the excess of information with minor importance. More specifically, in ANN models, the model inputs are the number of neurons in the input layer, the model parameters are the number of neurons and the values of interconnection weights, which do not have any physical meaning, in the hidden layers and, at last, the outputs are the number of neurons in the output layer [[34]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0170). I

In Lin et al. [[42]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0210) it is noted that the learning of long temporal dependencies with [gradient descent methods](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/gradient-descent-method) is more effective using a [recurrent](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/recurrent) architecture named Nonlinear Autoregressive with eXogenous input (NARX) [[43]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0215) than with the use of other models.

In the one-step-ahead prediction task, the ANN must estimate the next time series state, without giving the feedback to the model input [regressor](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/regressors). Namely, the input regressor contains only real data points from the time series. To achieve a larger prediction horizon, which is known as multistep-ahead prediction, the output of the model should be given as feedback to the input regressor in order to obtain a recursive prediction for a number of time steps in the future. In this case, the components of the input regressor, previously composed by actual time series values, are progressively replaced by predicted values forming a recurrent architecture.

The methods for feature selection can be classified into two categories: these ones that are based on model testing and those that directly apply on the data without the need to build and test models. Among the methods based on model testing, the principal ones are:

* •  
  Stepwise development. Adding up or removing input variable step by step, and testing the model performance gradually.
* •  
  Ad hoc. Developing models with different input variables sets, and comparing the performance.
* •  
  Global methods. [Genetic algorithms](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/genetic-algorithm) are an example.

Analytical methods. A statistical method measures the [strength](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/materials-science/mechanical-strength) in the relationship between the input and output variables. The most used method is the correlation which only takes into account the [linear dependence](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/linear-dependence) between variables. Another statistical methods, like the [mutual information](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/mutual-information), are necessary when nonlinear dependence is suspicious to be present.

In this work, it has been used analytical methods to measure the linear and nonlinear dependence by means of the correlation and mutual information, respectively. In addition, scatter-plots and model tests have been used in a complementary way.

A common technique to obtain an optimum embedding delay *τ* is to use the first [local minimum](https://www-sciencedirect-com.universidadviu.idm.oclc.org/topics/engineering/local-minimum) of the average mutual information function [[54]](https://www-sciencedirect-com.universidadviu.idm.oclc.org/science/article/pii/S0378778814005349#bib0270). The average mutual information determines how much information has in common the measure *x*(*t*) at a given time with another instant *x*(*t* + *t*).

10.- PREDICTION MODEL FOR THE ENERGY INDUSTRY: THE CASE OF SPAIN

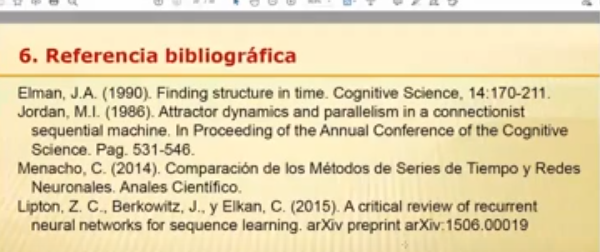
<https://www.proquest.com/docview/1682227745?parentSessionId=eDfkB2bmAiIUas9rO%2FS91YGVVshU9OltSQo17KON5fs%3D&pq-origsite=summon&accountid=198016&sourcetype=Scholarly%20Journals>

Most forecasting models are based on assumptions, in many cases certainly reliable and great value, but in which no energy variables jointly involved, political, economic and environmental, as done in this work. The Spanish energy model is very specific and requires an independent study for other countries due to the large energy dependence has Spain.[9,10]

11.- Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting

rtificial neural networks (ANNs) [3], support vector machines (SVMs) [4,5], and regression trees [6] have been applied successfully for diverse power demand prediction tasks. More recently, deep learning (DL) has emerged as a very powerful approach for time series forecasting. DL models are especially suitable for big-data temporal sequences due to their capacity to extract complex patterns automatically without feature extraction preprocessing steps [7]. As an evolution from simple ANNs, deep, fully connected networks have been applied for load forecasting problems [8]. However, fully connected networks are unable to capture the temporal dependencies of a time series. Consequently, more specialised DL models such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) started to gain importance in the time series forecasting field. These networks can efficiently encode the underlying patterns of time series by transforming the temporal problem into a spatial architecture [9].

<https://www.proquest.com/docview/2385521113?parentSessionId=5Jt6IOLZfhqM%2Fa%2FqHAbWV5rmZuEsLTmA4zvTX9IMgJY%3D&pq-origsite=summon&accountid=198016&sourcetype=Scholarly%20Journals>



<https://www.youtube.com/watch?v=MlktVhReO0E>

<https://www.youtube.com/watch?v=vJ6F5pO5uIY>

12.-

https://ebookcentral.proquest.com/lib/universidadviu/reader.action?docID=6420045

Energy demand forecasting is a type of demand forecasting where the goal is to predict the future load (or energy demand) on an energy grid. It is a critical business operation for companies in the energy sector, as operators need to maintain the fine balance between the energy consumed on a grid and the energy supplied to it. Typically, grid operators can take short-term decisions to manage energy supply to the grid and keep the load in balance. An accurate short-term forecast of energy demand is therefore essential for the operator to make these decisions with confidence. This scenario details the construction of a machine learning energy demand forecasting solution.

By using the statsmodels Python module, which has a tsa (time series analysis) package as well as the seasonal \_ decompose() function, we can visualize these for components and gain additional insights from our time series data.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 00:22:59.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 00:09:48.

algorithms when your time series data has input values with differing dimensions. Standardization assumes that your observations fit a Gaussian distribution (bell curve) with a well-behaved mean and standard deviation. This includes algorithms like support vector machines and linear and logistic regression and other algorithms that assume or have improved performance with Gaussian data.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 00:35:27.

The pandas library provides the shift() function to help create these shifted or lag features from a time series data set: this function shifts an index by desired number of periods with an optional time frequency. The shift method accepts a freq argument which can accept a DateOffset class or a timedelta -like object or also an offset alias

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 00:47:56.

Depending on your time series scenario, you can expand the window width and include more lagged features. A common question that data scientists ask before performing the operation of adding lag features is how large to make the window. A good approach would be to build a series of different window widths and alternatively add and remove them from the data set to see which one has a more evident positive effect on your model performance.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 00:51:28.

Due to their applicability to many real-life problems, such as fraud detection, spam email filtering, finance, and medical diagnosis, and their ability to produce actionable results, deep learning neural networks have gained a lot of attention in recent years. Generally, deep learning methods have been developed and applied to univariate time series forecasting scenarios, where the time series consists of single observations recorded sequentially over equal time increments (Lazzeri 2020). For this reason, they have often performed worse than naïve and classical forecasting methods, such as autoregressive integrated moving average (ARIMA). This has led to a general misconception that deep learning models are inefficient in time series forecasting scenarios, and many data scientists wonder whether it’s really necessary to add another class of methods, like convolutional neural networks or recurrent neural networks, to their time-series toolkit

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 01:19:17.

Reasons to Add Deep Learning to Your Time Series Toolkit – Deep learning neural networks are able to automatically learn arbitrary complex mappings from inputs to outputs and support multiple inputs and outputs. These are powerful features that offer a lot of promise for time series forecasting, particularly on problems with complex-nonlinear dependencies, multivalent inputs, and multi-step forecasting. These features, along with the capabilities of more modern neural networks, may offer great promise, such as the automatic feature learning provided by convolutional neural networks and the native support for sequence data in recurrent neural networks. In this section, you will discover the promised capabilities of deep learning neural networks for time series forecasting. Specifically, we will discuss the following deep learning methods’ capabilities: ■■ ■■ ■■ ■■ Deep learning neural networks are capable of automatically learning and extracting features from raw and imperfect data. Deep learning supports multiple inputs and outputs Recurrent neural networks— specifically long short-term memory (LSTM) networks— are good at extracting patterns in input data that span over relatively long sequences.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 01:22:11.

Recurrent Neural Networks for Time Series Forecasting – In this section I will introduce a very popular type of artificial neural networks: recurrent neural networks, also known as RNNs. Recurrent neural networks are a class of neural networks that allow previous outputs to be used as inputs while having hidden states.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 01:22:29.

Deep learning is a subset of machine learning. Deep learning algorithms are specific types of machine learning algorithms that are based on artificial neural networks. In this case, the learning process is based on the same steps as those for machine learning, but it is called deep because the structure of algorithm is based on artificial neural networks that consist of multiple input, output, and hidden layers, containing units that transform the input data into an information that the next layer can use to perform a certain automated predictive task once the deep learning model is deployed.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 01:24:32.

Neural networks can be useful for time series forecasting problems by eliminating the immediate need for massive feature engineering processes, data scaling procedures, and making the data stationary by differencing.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 13:51:05.

On the contrary, neural networks are robust to noise in input data and in the mapping function and can even support learning and prediction in the presence of missing values. Convolutional neural networks (CNNs) are a category of neural networks that have proven very effective in areas such as image recognition and classification. CNNs have been successful in identifying faces, objects, and traffic signs, apart from powering vision in robots and self-driving cars. CNNs derive their name from the “convolution” operator (Lazzeri 2019a).

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 13:54:22.

Specifically, neural networks can be configured to support an arbitrary but fixed number of inputs and outputs in the mapping function. This means that neural networks can directly support multivariate inputs, providing direct support for multivariate forecasting. A univariate time series, as the name suggests, is a series with a single time-dependent variable. For example, we want to predict next energy consumption in a specific location: in a univariate time series scenario, our dataset will be based on two variables, time values and historical energy consumption observations (Lazzeri 2019a). A multivariate time series has more than one time-dependent variable. Each variable not only depends on its past values but also has some dependency on other variables. This dependency is used for forecasting future values. Let’s consider the above example again.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 14:04:30.

With neural networks, an arbitrary number of output values can be specified, offering direct support for more complex time series scenarios that require multivariate forecasting and even multi-step forecast methods. There are two main approaches that deep learning methods can be used to make multi-step forecasts: ■■ ■■ Direct , where a separate model is developed to forecast each forecast lead time Recursive , where a single model is developed to make one-step forecasts, and the model is used recursively where prior forecasts are used as input to forecast the subsequent lead time The recursive approach can make sense when forecasting a short contiguous block of lead times, whereas the direct approach may make more sense when forecasting discontiguous lead times.

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 14:08:13.

Recurrent neural networks (RNNs) were created in the 1980s but have just recently been gaining popularity due to increased computational power from graphic processing units. They are especially useful with sequential data because each neuron or unit can use its internal memory to maintain information about the previous input. An RNN has loops in it that allow information to be carried across neurons while reading in input. However, a simple recurrent network suffers from a fundamental problem of not being able to capture long-term dependencies in a sequence. This is a major reason RNNs faded out from practice until some great results were achieved with using a LSTM unit inside the neural network. Adding the LSTM to the network is like adding a memory unit that can remember context from the very beginning of the input (Lazzeri 2019a).

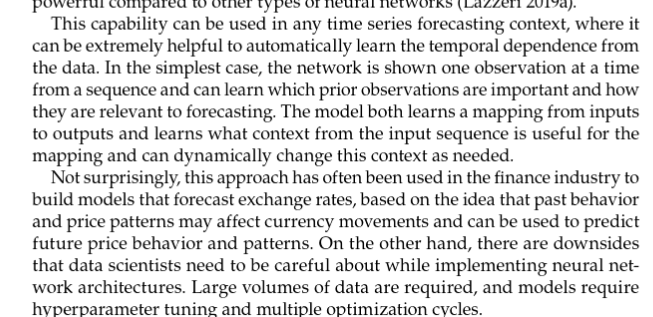
Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

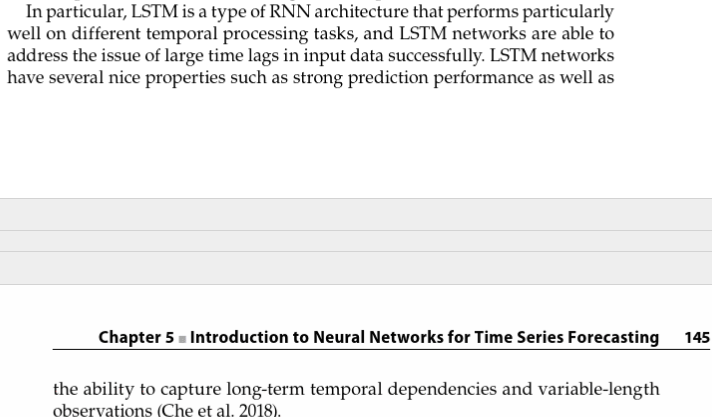
Created from universidadviu on 2024-10-21 14:56:35.

The example of video processing can be very effective when we need to understand how LSTM networks work: in a movie, what happens in the current frame is heavily dependent upon what was in the last frame. Over a period of time, an LSTM network tries to learn what to keep and how much to keep from the past and how much information to keep from the present state, which makes it powerful compared to other types of neural networks

Lazzeri, Francesca. *Machine Learning for Time Series Forecasting with Python*, John Wiley & Sons, Incorporated, 2020. *ProQuest Ebook Central*, http://ebookcentral.proquest.com/lib/universidadviu/detail.action?docID=6420045.

Created from universidadviu on 2024-10-21 14:58:04.





TODO DESDE LA PAG 145 SIRVE

13.- de este artículo puedo sacar la parte matemática de LSTM y GRU. También las ecuaciones de los errores.

Building Energy Time Series Data Mining for Behavior Analytics and Forecasting Energy consumption

<https://web-p-ebscohost-com.universidadviu.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=0&sid=6126dbd5-f241-4384-958a-55a0d487e027%40redis>

14.- Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting

<https://www.proquest.com/docview/2385521113?accountid=198016&parentSessionId=5Jt6IOLZfhqM%2Fa%2FqHAbWV5rmZuEsLTmA4zvTX9IMgJY%3D&parentSessionId=ScGZTQ2fzD4bpWDXObMxRGnTAALFwtj8grVGI5fARoQ%3D&pq-origsite=summon&sourcetype=Scholarly%20Journals>

espite the popularity of RNNs, several works using convolutional networks can be found. In both [17,18], the authors proposed CNN models for short-term load forecasting that provides comparable results to LSTM models. Other works have been able to build deep convolutional networks that can outperform LSTM networks for electricity demand [19] and solar power data problems [20]. Furthermore, in all these works, the CNN models proved to be more suitable for real-time applications given their faster training and testing execution time.

More recently, a specialised CNN architecture known as temporal convolutional networks (TCN) has acquired popularity due to their suitability to deal with time series data. TCNs were first proposed in [24], in which they were compared to several RNNs over sequence modelling tasks. TCNs use causal dilated causal convolution in order to be able to capture longer-term dependencies and prevent information loss. Furthermore, they present other advantages over RNNs such as lower memory requirements, parallel processing of long sequences as opposed to the sequential approach of RNNs, and a more stable training scheme

As suggested in recent forecasting studies that use neural networks [33,34], in this work, we adopt the MIMO strategy (Multi-Input Multi-Output) which belongs to the last category. Instead of forecasting each time-step independently, the MIMO approach can model the dependencies between the predicted values since it outputs the complete forecasting window. Furthermore, this strategy avoids the accumulated errors over predictions that appear in the recursive strategy.

Following this approach, a moving window scheme is used to create the input–output pairs that will be fed to the neural network. All deep learning models used in this study accept a fixed-length window as input and have an output dense layer with as many neurons as the forecasting horizon defined for each problem (24 for electricity demand and 48 for electric vehicle demand). Figure 3 illustrates the process of applying the moving window over the complete time series. As can be seen, the window slides and obtains an input–output instance at each position. While the output window size is defined by the problem, the input window size has to be decided. The optimal value can be different depending on the data, the designed model, and the forecasting horizon. In our study, we have experimented with three different sizes for the input window of each problem. The values have been carefully selected, considering the characteristics and seasonality of the datasets. For the electricity demand, we evaluate using 144, 168, and 268 time-steps as input window (which corresponds to 24, 28, and 48 h, respectively)

IDEAS  
- HAcer gráficos de las regiones que más consumen energía.

* Hacer un gráfico de las regiones que más tienen temperaturas extremas
* Hacer un gráfico de correlación de la energía de las regiones y las temperaturas.
* Hablar de en qué épocas del año hay más demanda de energía.
* Separar los datos o poner una columna con el nombre de la temporada (verano, otoño, invierno, primavera)
* De la serie de la demanda de energía ver si es una serie estacionaria o no. (se puede hacer una prueba de dicky fuller) Sacar estadísticas descriptivas.
* Discusión: pone que se agregue el precio, también la tasa de desempleo,
* Los modelos estadísticos clásicos parten de que las relaciones de las variables son lineales. Mientras que las redes neuronales parten de que las relaciones son no lineales.
* ver el consumo de enegería en las diferente épocas del año.
* Añadir más variables al dataset con el método de la ventana explicado en el libro.
* Scalar.. y luego deshacer el scalado