

# DEEP LEARNING EN LA PREDICCIÓN DE GENERACIÓN DE UN PARQUE EÓLICO

J.M. Torres<sup>1</sup>, R.M. Aguilar<sup>2</sup>, J.A. Méndez<sup>3</sup> y K.V. Zúñiga-Meneses<sup>4</sup>  
Dpto. Ingeniería Informática y de Sistemas, Universidad de La Laguna  
38200 La Laguna (Tenerife), España.

<sup>1</sup>jmtorres@ull.edu.es, <sup>2</sup>raguilar@ull.edu.es, <sup>3</sup>jamendez@ull.edu.es y <sup>4</sup>zm.victoriak@gmail.com

## Resumen

*Uno de los grandes retos tecnológicos actuales es la incorporación de las energías renovables al sistema eléctrico. El objetivo es conseguir que la generación eléctrica sea sostenible y respetuosa con el medioambiente, así como abordable económicamente. Sin embargo para que esta incorporación tenga éxito es necesario disponer de herramientas de predicción que permitan conocer con suficiente antelación la cantidad de energía de origen renovable que estará disponible para ser inyectada en la red; permitiendo ajustar adecuadamente el resto de fuentes de generación con el objeto de suplir la demanda, incluidas las basadas en combustibles fósiles. Esto permitiría limitar el impacto ambiental y la dependencia con respecto este tipo de carburantes en un previsible escenario de escasez.*

*En este trabajo se quiere avanzar en la creación de dichos modelos de predicción de la generación de los parques eólicos utilizando aprendizaje profundo o deep learning. En este artículo se presenta un modelo de predicción basado en una red neuronal profunda multicapa que, a partir de la predicción de las condiciones atmosféricas, es capaz de estimar 24 horas antes la generación producida por un parque eólico situado en la isla de Tenerife.*

**Palabras clave:** deep learning, aprendizaje profundo, predicción, generación eólica.

## 1. INTRODUCCIÓN

Uno de los grandes retos actuales es conseguir que la generación eléctrica sea sostenible y que la distribución de energía sea respetuosa con el medioambiente, abordable económicamente y aceptable socialmente. Obviamente el objetivo principal es la transición hacia un sistema energético que permita reducir la dependencia respecto a los carburantes fósiles, en un escenario en el que se contempla simultáneamente: la escasez de los mismos, el crecimiento de la demanda energética a nivel mundial y el grave impacto que tiene el uso de este tipo de combustibles en el medioambiente.

Por todo esto se considera fundamental la incorpo-

ración de las energías renovables al sistema eléctrico. Sin embargo, para que esta incorporación tenga éxito es necesario disponer de herramientas de predicción que permitan conocer con suficiente antelación la cantidad de energía de origen renovable que estará disponible para ser inyectada en la red en un momento dado. Una predicción adecuada de la generación de energías renovables permite planificar con antelación las necesidades de otras fuentes de generación, con el objetivo de asegurar que se suple la demanda energética al tiempo que se minimiza el uso de fuentes de energía basadas en combustibles fósiles.

En general los sistemas de redes eléctricas se pueden clasificar en dos tipos: continentales y aislados[18]. En los primeros se puede incrementar la generación mediante renovables de forma bastante robusta, debido a que la interconexión de sistemas eléctricos puede manejar la impredecibilidad de las fuentes de energía renovables. Sin embargo los sistemas eléctricos aislados —como los que tienen muchas islas— presentan muchas dificultades para incrementar la generación mediante renovables manteniendo las debidas garantías en el suministro. Por ejemplo, en un sistema eléctrico aislado donde la energía eólica o fotovoltaica tuviera mucha importancia, las variaciones no previstas en la generación, debido a cambios en la velocidad del viento o en la radiación solar, pueden generar un desequilibrio entre la energía generada y la demandada sin que hayan conexiones con otros sistemas eléctricos que puedan ayudar a compensarla.

En los sistemas aislados es fundamental coordinar las diferentes fuentes de energía para garantizar el suministro. Por ejemplo, en los valles de la generación eólica es necesario utilizar otras fuentes de energía para compensarla. Fuentes que pueden ser renovables o basadas en combustibles fósiles, según el caso. Sea como fuere, es evidente que para que esta coordinación sea posible es fundamental poder predecir con precisión la generación de las fuentes de energía renovable, para planificar las necesidades del sistema con antelación.

Lamentablemente la energía eólica, que es una de las fuentes de energía más utilizadas a nivel mun-

dial, tiene como principal desventaja lo difícil que es hacer predicciones precisas; debido a su fuerte dependencia de las condiciones atmosféricas y lo difíciles que estas son de predecir. Por ejemplo, en la isla de Tenerife el Instituto Tecnológico de Energías Renovables (ITER) es la entidad que gestiona el mayor parque eólico. Sus predicciones de generación se estiman mediante regresión polinomial a partir de la velocidad del viento prevista según la predicción meteorológica para la ubicación del parque. Las predicciones de generación son entregadas al operador de la red con 24 horas de antelación y presentan un margen de error de entre el 20 y el 60 %, lo que se traduce en importantes pérdidas debido a las multas impuestas por el operador. A fin de cuentas en muchos casos los errores de predicción se traducen en el aumento de generación con plantas convencionales.

En este artículo presentaremos un modelo de predicción basado en una red neuronal profunda multicapa que, a partir de la predicción de las condiciones atmosféricas, es capaz de estimar 24 horas antes la generación producida por un parque eólico situado en la isla de Tenerife. El objetivo es mejorar la previsión de la generación de dicho parque eólico respecto a la técnica de regresión usada actualmente.

### 1.1. ESTADO DEL ARTE

La relación entre la potencia eléctrica generada por una turbina y la velocidad del viento viene indicada por la curva de potencia del generador, que debe proporcionar el fabricante de la misma. En la figura 1 se puede observar un ejemplo.

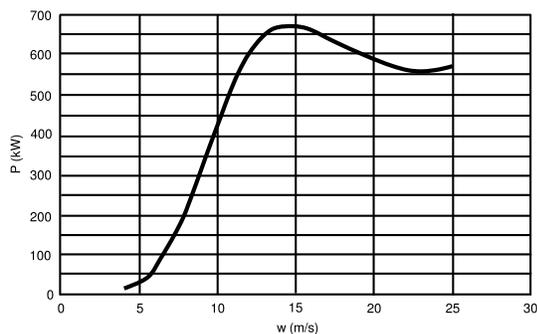


Figura 1: Curva de potencia del generador MADE AE-46[1].

Sin embargo esto es una simplificación porque realmente, si bien la potencia eléctrica  $P$  generada por una turbina viene determinada fundamentalmente por la velocidad del viento  $w$ , también influyen el área  $A$  barrida por el generador y la densidad del aire  $\rho$ [16],

$$P = 0,5\rho Aw^3 \quad (1)$$

que a su vez se ve afectada por la humedad relativa del aire y la temperatura ambiente. Por lo que resulta evidente la relación entre la potencia eléctrica generada y las condiciones meteorológicas.

Actualmente se pueden utilizar varias técnicas para la predicción de la generación eólica a corto plazo. Las más sencillas utilizan el histórico de potencia generada  $P$ , pero por lo general el punto de partida es la predicción de las condiciones meteorológicas. A partir de ellas las distintas técnicas difieren en cómo se utilizan para convertirlas en predicciones de la generación eólica. En general existen dos aproximaciones a este problema, las basadas en modelos físicos y las basadas en modelos estadísticos[15].

Las técnicas basadas en modelos físicos consisten en simular las variaciones del flujo del viento a lo largo del parque con el objeto de estimar su magnitud y orientación en cada generador. Eso permite utilizar después las curvas de potencia proporcionadas por el fabricante para estimar la generación total del parque. Por el contrario la aproximación mediante modelos estadísticos no requiere disponer de un modelo físico del parque. Simplemente se utiliza un modelo estadístico para predecir directamente la potencia generada a partir de las condiciones meteorológicas. Los parámetros del modelo estadístico son estimados ajustando el modelo a datos históricos de generación y de las condiciones meteorológicas. En algunos casos una combinación de ambas aproximaciones puede ser necesaria para obtener predicciones lo suficientemente precisas. En esos casos el modelo físico se utiliza para obtener la mejor estimación de la velocidad y la dirección del viento local en cada generador, usándose posteriormente un modelo estadístico para reducir aun más el error de la predicción.

#### 1.1.1. Predicción a 4 o 6 horas

En la mayor parte de los casos las aproximaciones estadísticas hacen uso del análisis de serie temporales, si bien no son la única alternativa. El estimador de referencia con respecto al que se compara se basa en considerar que la generación futura será igual que la actual, medida como el promedio en una ventana de tiempo anterior. Pese a su sencillez, es muy complicado obtener mejores resultados que el estimador de referencia en horizontes de minutos a 4 o 6 horas.

Diversos autores han hecho propuestas que mejoran los resultados del estimador de referencia. Por

ejemplo, Bossany et al.[4] utiliza el filtro de Kalman para obtener una mejora del 10 % en el error RMS de la predicción para el siguiente instante de tiempo respecto al estimador de referencia con una ventana de tiempo de 1 min. Esta mejora se pierde cuando el estimador de referencia utiliza promedios de 1 hora. Tantareanu et al.[23] estima que los modelos ARMA, aplicados directamente sobre los datos de generación eólica, pueden alcanzar un rendimiento hasta un 30 % mejor para de 3 a 10 instantes de tiempo posteriores respecto al estimador de referencia en promedios de 4 seg. con datos muestreados a 2.5Hz. Dutton et al.[5] usó un modelo lineal autoregresivo y un modelo de lógica difusa adaptativo que ofreció mejoras de hasta el 20 % en el error RMS para un horizonte de predicción de 8 horas. Sin embargo, para horizontes mayores la banda de confianza 95 % contenía la mayor parte de los valores de velocidad del viento probables, por lo que se consideró que utilizar un enfoque basado en la predicción meteorológica era más prometedor.

En general los modelos de series temporales como ARIMA o Box-Jenkins[19], que suelen ser útiles en tareas de predicción en determinados procesos industriales, proporcionan resultados razonablemente buenos para horizontes de hasta 6 horas en la predicción de energía eólica utilizando solamente los datos previos de generación. Existen otros autores que han intentado utilizar modelos con técnicas más próximas a la inteligencia artificial, como es el caso de los modelos basados en redes neuronales. Sin embargo, los pocos beneficios obtenidos respecto al estimador de referencia no parecen ser suficientes para compensar el esfuerzo extra de entrenar dichos modelos.

### 1.1.2. Predicción a más de 6 horas

Ninguno de los modelos anteriores resulta útil para predicciones a horizontes de entre 24 y 48 horas, que son los que suscitan más interés para la operación del sistema y la participación en el mercado energético. Para esos horizontes es preferible obtener predicciones de las condiciones meteorológicas que posteriormente son utilizadas como entrada para el modelo de predicción de la generación eólica. Estos modelos pueden estar basados en series temporales, pero también se ha tenido éxito utilizando técnicas más próximas a la inteligencia artificial, como por ejemplo modelos basados en redes neuronales o regresión de soporte vectorial (SVR). En este sentido distintos autores han probado diferentes aproximaciones de las cuales las basadas en redes neuro-fuzzy parecen ser las que mejor rendimiento[6, 11, 12] ofrecen.

Es destacable el sistema de ARMINES[7] que in-

tegra modelos basados en series temporales con modelos basados en neuro-fuzzy. Los modelos basados en series temporales permiten predecir la generación en horizontes de hasta 10 horas, pues el componente autoregresivo mejora significativamente la predicción en periodos en los que el uso exclusivo de la estimación de las condiciones meteorológicas no es suficiente. Mientras los modelos basados en neuro-fuzzy permiten alcanzar horizontes de hasta 72 horas. Un sistema inteligente pondera ambos tipos de predicción para obtener el máximo rendimiento en todo el intervalo, con un error promedio de aproximadamente el 10 % de la potencia instalada.

Otro ejemplo es el del sistema WPMS que proporciona predicciones para el 95 % de la potencia instalada en Alemania. En este caso la predicción meteorológica sirve como entrada a un red neuronal artificial que se encarga de estimar la potencia generada[20].

### 1.1.3. Aprendizaje profundo

En este trabajo se quiere avanzar en la creación de modelos de predicción de la generación de los parques eólicos utilizando aprendizaje profundo o *deep learning* en lugar de las técnicas utilizadas actualmente.

El aprendizaje profundo comprende un conjunto de técnicas y algoritmos que facilitan crear redes neuronales con un número importante de capas y neuronas. Una de las diferencias clave del aprendizaje profundo respecto a otras técnicas de aprendizaje automático es que no debe ser necesario seleccionar y preparar las mejores características de los datos que van a ser utilizadas como entrada al modelo —la llamada ingeniería de características—. Si no que la propia red, a través de sus múltiples capas, es capaz de extraer por sí misma características cada vez más abstractas a partir de los propios datos usados durante el entrenamiento.

El uso de técnicas basadas en redes neuronales con un gran número de capas y parámetros presenta tres problemas fundamentales. El primero es la necesidad de contar con un gran número de muestras de entrenamiento, que deben ser manejadas por las arquitecturas actuales pero que, sobre todo, hacen imposible utilizar el clásico aprendizaje en lote —donde en cada paso del entrenamiento se consideran todas las muestras al mismo tiempo—. El segundo es el importante costo computacional de entrenar este tipo de redes. Y el tercero es que durante el entrenamiento la propagación hacia atrás de los errores sufre un fenómeno denominado desvanecimiento del gradiente[9] que dificulta notablemente el entrenamiento de toda la red.

Para poder manejar grandes conjuntos de muestras se acabaron imponiendo las técnicas de entrenamiento basadas en el aprendizaje online. En concreto una variante del gradiente descendiente estocástico que usa minilotes de muestras seleccionadas aleatoriamente[22]. Algunos trabajos[10, 3] relativos a la correcta inicialización de los pesos ayudaron a resolver el problema del desvanecimiento del gradiente. Además surgieron nuevas ideas como: las funciones de activación ReLU[8], la técnica de regulación *dropout*[21], mecanismos para detener el entrenamiento cuando no se espera que vayan a seguir mejorando los resultados de forma significativa, propuestas para seleccionar una tasa de aprendizaje adecuada o incluso para no tener que escoger ninguna y optimizadores más eficientes. Con todo ello el problema de conseguir una inicialización adecuada del modelo se hizo menos importante, haciendo posible utilizar alternativas más simples

El aumento del tamaño de las redes también puso de relieve las importantes necesidades computacionales de este tipo de soluciones. Por lo tanto se hizo necesario aprovechar la paralelización que ofrecen las arquitecturas multi-procesador y multi-núcleo y, sobre todo, las tremendas posibilidades de las GPU modernas. Debido a las dificultades para desarrollar técnicas y algoritmos e implementarlas para arquitecturas tan diferentes como las anteriores, comenzaron a surgir distintas bibliotecas y entornos que facilitan el desarrollo sin tener que conocer los detalles de implementación en la arquitectura del hardware subyacente. Entre estas librerías cabe destacar Caffe, Torch, Theano y Tensorflow[2]; siendo esta última la que empleamos para este trabajo. En la figura 2 se puede observar un ejemplo parcial de una red neuronal MLP implementada en Tensorflow.

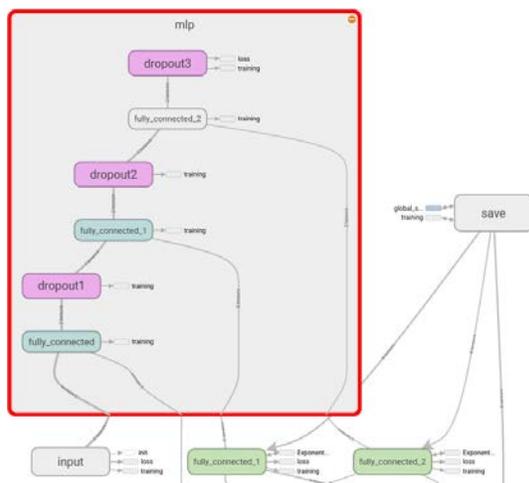


Figura 2: Modelo parcial MLP con Tensorflow.

En la actualidad la mayor parte de las aplicaciones

del aprendizaje profundo se centran en el reconocimiento del habla, la visión artificial y el procesamiento natural del lenguaje. Esto en parte es debido al desarrollo de las capas de convolución[17] que son especialmente eficientes extrayendo características de alto nivel de abstracción a partir de las entradas, de forma similar a como ocurre en el cortex visual[14]. Afortunadamente esta capacidad de las redes neuronales profundas para extraer características cada vez más abstractas también tiene mucha utilidad en problemas de regresión más simples.

Por ejemplo, en el problema de la predicción de la generación eólica se dispone de los resultados de la predicción de las condiciones meteorológicas. Esta predicción puede contener errores que una red neuronal profunda podría aprender a predecir y corregir, a través del análisis de los datos de entrada. De forma similar, la relación entre las condiciones meteorológicas y la generación eólica no es lineal. Las curvas de potencia teóricas de los fabricantes de los generadores no siempre coinciden con las reales y además pueden variar con el tiempo. Finalmente, en la relación entre las condiciones meteorológica y la generación total del parque sin duda influyen otros factores, como la disposición de los generadores o la orografía del terreno. Por lo tanto, la hipótesis de partida es que las redes neuronales profundas pueden aprender estas relaciones abstractas y usarlas para obtener una mejor predicción de la generación eólica.

## 2. PROBLEMA

El Instituto Tecnológico y de Energías Renovables (ITER) gestiona tres parques eólicos. Entre ellos se encuentra el parque MADE, que cuenta con una potencia nominal de 4800 kW gracias a 8 generadores MADE AE-46. Dos veces al día se genera una predicción meteorológica de las siguientes 48 horas en periodos de una hora. Una vez al día se toma la velocidad del viento prevista para entre las próximas 24 y 48 horas y se hace una regresión polinomial con la que se estima la generación para cada hora de dicho intervalo. Esta predicción se remite al operador de la red.

El objetivo de este trabajo es mejorar dicha predicción utilizando redes neuronales profundas en lugar de la regresión polinomial que se utiliza actualmente. En la entrada del nuevo modelo se inyectan una serie de variables relativas a la predicción de las condiciones meteorológicas, dando como salida la predicción de la generación para todo el parque.

### 2.0.1. Entradas

A partir de la ecuación (1) se decidió usar como entrada del modelo la velocidad del viento  $w$  prevista. También se incluyeron la humedad relativa y la temperatura, por su relación con la densidad del aire  $\rho$ . Por el contrario no consideramos el área barrida por el generador  $A$  por estimarla constante.

Aunque en la ecuación (1) no se indica, es evidente que la dirección del viento también está muy relacionada con la potencia eléctrica generada, por lo que igualmente se incluyó entre las entradas. De la misma manera se incluyó la presión atmosférica, el día del año y la hora del día, con el objeto de que el modelo tuviera información para contemplar los ciclos locales de vientos diarios.

Finalmente se incluyó la fracción de generadores fuera de servicio en cada hora, puesto que algunas de las muestras de la generación total del parque, contra las que se ajustó la salida del modelo, fueron obtenidas cuando algunos generadores estaban fuera de servicio. En la tabla 1 se resumen las entradas del modelo.

Tabla 1: Resumen de las entradas.

Número	Descripción
1	Día del año
2	Hora del día
3	Velocidad del viento (m/s)
4	Dirección del viento (°)
5	Humedad relativa
6	Temperatura (°C)
7	Presión atmosférica (Pa.)
8	Generadores fuera de servicio

### 2.0.2. Redes neuronales

Los modelos utilizados fueron redes neuronales perceptrón multicapa (MLP) completamente conectadas con función de activación ReLU. En las redes MLP completamente conectadas todas las neuronas de una capa reciben como entrada todas las salidas de las neuronas de la capa anterior. En nuestro caso concreto la función de activación tanto de la capa de entrada como de las capas ocultas es ReLU[8], mientras que la capa de salida es lineal. La función de activación ReLU es muy práctica en problemas de regresión porque no está limitada entre  $[0, 1]$ , favorece la dispersidad de la solución en las neuronas ocultas y no sufre el problema del desvanecimiento del gradiente comentado anteriormente.

Para evitar sobreajustar los modelos al entrenarlos, se introdujo en la función de costo la regularización de los parámetros mediante la norma L2.

También, en algunos casos concretos, se hicieron pruebas con *dropout* para comprobar sus efectos al tratar de generalizar más los modelos entrenados.

Para el entrenamiento se dispuso de datos muestreados cada hora, desde enero de 2014 a abril de 2014, que se separaron aleatoriamente en tres conjuntos. Un 60% conformó el conjunto de datos de entrenamiento, un 20% el de datos de validación y el 20% restante el de datos de prueba. Para componer los conjuntos de validación y prueba sólo se consideraron aquellas muestras tomadas cuando todos los generadores estuvieron en servicio durante el periodo medido. Tanto los datos de entrada como de salida de las muestras se estandarizaron estadísticamente con el objeto de conseguir la máxima eficiencia en el entrenamiento.

El aprendizaje se realizó con el optimizador Adam[13], debido a que este algoritmo no necesita que se le indique una tasa de aprendizaje, usando minilotes[22] de 128 muestras escogidas aleatoriamente. Se validó el modelo ajustado cada 10 iteraciones completas sobre el conjunto de datos de entrenamiento, utilizando como criterio de parada el que el MAE de evaluación decreciera menos de  $10^{-3}$ .

## 3. RESULTADOS Y CONCLUSIONES

El error medio de predicción de la regresión polinomial que utiliza actualmente el parque es del 11% de la potencia instalada. Mientras que los mejores resultados obtenidos utilizando aprendizaje profundo ofrecen un error medio del 10%. Además, gracias a la batería de pruebas realizada con diferentes diseños de red neuronal profunda, se aprecia que tiene mucho más efecto en la reducción del error aumentar el número de neuronas por capa, en alguna de las capas ocultas, que aumentar el número de capas.

Por lo tanto, no podemos afirmar que las redes neuronales MLP sean capaces de extraer algún tipo de característica oculta en los datos ni de ofrecer una mejora sustancial en la predicción de la generación eólica. Sin embargo es importante tener en cuenta que en el caso de las series temporales es el componente autoregresivo el que permite extender el horizonte de predicción más allá de lo que se puede hacer usando sólo la predicción meteorológica. Queda por tanto como línea abierta analizar otras arquitecturas de aprendizaje profundo para corroborar que ocurre o no lo mismo. Nos referimos a redes que admiten introducir lotes de series temporales de muestras o a las que entran dentro de la categoría de las redes neuronales

recurrentes.

### Agradecimientos

Queremos agradecer al Instituto Tecnológico y de Energías Renovables (ITER), especialmente a Guillermo Galván García, Jefe del Departamento de Eólica; y a Jonatán Felipe García, Gestor de proyectos I+D del Área de Ingeniería, por su colaboración.

### Referencias

- [1] Fabricantes de aerogeneradores y características de sus máquinas. *Energía: Ingeniería energética y medioambiental*, 29(174):150–196, 2003.
- [2] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. {TensorFlow}: Large-Scale Machine Learning on Heterogeneous Systems, 2015.
- [3] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. Greedy Layer-Wise Training of Deep Networks. In Bernhard Schölkopf, John Platt, and Thomas Hoffman, editors, *Advances in Neural Information Processing Systems 19 (NIPS'06)*, pages 153–160. MIT Press, 2007.
- [4] E A Bossanyi. Short-Term wind prediction using Kalman Filters. *Wind Engineering*, 9(1):1–8, 1985.
- [5] A G Dutton, Georges Kariniotakis, Jackie A. Halliday, and Eric Nogaret. Load and Wind power Forecasting Methods for the optimal management of Isolated Power Systems with High Wind Penetration. 23(2):69–87, 1999.
- [6] H Fukuda and others. The development of a wind velocity prediction method based on a data-Mining type auto-regressive model. In *Proceedings of the European Wind Energy Conference*, pages 741–744, Copenhagen, Denmark, 2001.
- [7] Gregor Giebel and Et. Al. The State of the Art in Short-Term Prediction of Wind Power. Technical report, 2011.
- [8] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep Sparse Rectifier Neural Networks. In Geoffrey J Gordon and David B Dunson, editors, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS-11)*, volume 15, pages 315–323. Journal of Machine Learning Research - Workshop and Conference Proceedings, 2011.
- [9] Xavier Glorot, Xavier Glorot, and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10)*. Society for Artificial Intelligence and Statistics, 2010.
- [10] G E Hinton. Reducing the Dimensionality of Data with Neural Networks. *Science*, 313(5786):504–507, 7 2006.
- [11] Georges Kariniotakis. *Contribution au développement d'un système de contrôle avancé pour les systèmes éolien-diesel autonomes*. PhD thesis, Ecole des Mines de Paris, Spécialité énergétique, 1996.
- [12] G.N. Kariniotakis, G.S. Stavrakakis, and E.F. Nogaret. Wind power forecasting using advanced neural networks models. *IEEE Transactions on Energy Conversion*, 11(4):762–767, 1996.
- [13] Diederik Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. 12 2014.
- [14] Norbert Krüger, Peter Janssen, Sinan Kalkan, Markus Lappe, Ales Leonardis, Justus Piater, Antonio J Rodríguez-Sánchez, and Laurenz Wiskott. Deep hierarchies in the primate visual cortex: what can we learn for computer vision? *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1847–71, 8 2013.
- [15] Lars Landberg, Gregor Giebel, Henrik Aalborg Nielsen, Torben Nielsen, and Henrik Madsen. Short-term Prediction—An Overview. *Wind Energy*, 6(3):273–280, 7 2003.
- [16] J. F. Manwell, J. G. McGowan, and Anthony L. Rogers. *Wind energy explained : theory, design and application*. Wiley, 2009.
- [17] Masakazu Matsugu, Katsuhiko Mori, Yusuke Mitari, and Yuji Kaneda. Subject independent facial expression recognition with ro-

- bust face detection using a convolutional neural network. *Neural Networks*, 16(5):555–559, 2003.
- [18] S. Padrón, J.F. Medina, and A. Rodríguez. Analysis of a pumped storage system to increase the penetration level of renewable energy in isolated power systems. Gran Canaria: A case study. *Energy*, 36(12):6753–6762, 12 2011.
- [19] Alan Pankratz, editor. *Forecasting with Univariate Box-Jenkins Models*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, USA, 8 1983.
- [20] K. Rohrig, K. Rohrig, B. Ernst, M. Hoppe-Kilpper, and et al. *2nd World Wind Energy Conference 2003. Proceedings. CD-ROM*. 2003.
- [21] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.
- [22] Ilya Sutskever, James Martens, George E Dahl, and Geoffrey E Hinton. On the importance of initialization and momentum in deep learning. In *ICML (3)*, volume 28 of *JMLR Proceedings*, pages 1139–1147. JMLR.org, 2013.
- [23] C Tantareanu and Nordvestjysk Folkecenter for Vedvarende Energi. *Wind Prediction in Short Term: A First Step for a Better Wind Turbine Control*. Nordvestjysk Folkecenter for Vedvarende Energi, 1992.