

OPENCOURSEWARE
REDES DE NEURONAS ARTIFICIALES
Javier Huertas Tato



Descripción Dominio 1: Solar Oklahoma

Este dominio de datos es un problema simplificado del planteado en una competición en el marco de predicción de energía solar [1] (<https://www.kaggle.com/c/ams-2014-solar-energy-prediction-contest>). Los datos para la realización de los casos prácticos se pueden encontrar como *DatosPracticaSolarOklahoma.csv* [PR-F-02.2].

El objetivo es predecir la energía solar diaria en 98 estaciones de monitorización ubicadas en el estado de Oklahoma. Para la labor de predicción se utilizan variables meteorológicas recogidas desde una malla de puntos proporcionadas por el sistema numérico de predicción meteorológica GEFS.

La simplificación del conjunto de datos consiste en utilizar sólo una estación de monitorización y cuatro puntos del GEFS (Global Ensemble Forecast System).

Para cada punto de la rejilla simplificada existen las siguientes variables meteorológicas:

- *dlwrf_s*: Flujo de radiación media descendente de onda larga en la superficie (W m⁻²)
- *uswrf_s*: Radiación media ascendente de onda corta en la superficie (W m⁻²)
- *dswrf_s*: Flujo de radiación media descendente de onda corta en la superficie (W m⁻²)
- *ulwrf_t*: Radiación media ascendente de onda larga sobre la atmósfera (W m⁻²)
- *pwat_ea*: precipitaciones sobre la profundidad de la atmósfera (kg m⁻²)
- *pres_ms*: Presión a nivel del mar (Pa)
- *spfh_2m*: Humedad específica a 2m sobre el suelo (kg)
- *ulwrf_s*: Radiación media ascendente de onda larga en la superficie (W m⁻²)
- *tmp_sfc*: Temperatura superficial (K)
- *tmin_2m*: Temperatura mínima sobre las pasadas 3h a 2m sobre el suelo (K)

Estas variables se en 5 instantes distintos de tiempo (a las 12, 15, 18, 21 y 23 horas). Por tanto usando 10 variables meteorológicas, 5 instantes de tiempo y 4 estaciones, el conjunto de datos tiene 200 variables de entrada y 1 variable objetivo (Energía media en W m⁻² del día en una estación).

En este conjunto de datos se dispone de días pertenecientes a 2005, 2006, 2007 y 2008, en total suman 1461 días.

[1] A. McGovern , D.J. Gagne , J. Basara , T.M. Hamill , D. Margolin , Solar energy prediction: an international contest to initiate interdisciplinary research on compelling meteorological problems, *Bull. Am. Meteorol. Soc.* 96 (8) (2015) 1388–1395.