

Información auxiliar sobre el *dataset*

1. Estructura de archivos

- ***column.spec.pdf***: este documento, contiene información para facilitar interpretar el dataset adjunto
- ***complimentary_info.pdf***: documento con información auxiliar, pero no imprescindible, para guiar el trabajo de desarrollo e interpretación de los algoritmos genéticos. Puede ayudar a plantear el frente de Pareto llegado el momento
- ***combined_model.io.csv***: CSV con el dataset completo, sin separación entre viento lento y viento rápido. En principio, los datos deberían ser demasiado ruidosos como para aportar información relevante, pero puede merecer la pena explorar su uso directo como última fase del TFM si el tiempo y las fuerzas lo permiten
- ***slow_model.io.csv***: CSV con datos únicamente referidos a predicciones de viento lento
- ***fast_model.io.csv***: CSV con datos únicamente referidos a predicciones de viento rápido

La separación entre viento lento y viento rápido se ha realizado utilizando la columna *forecast*. El criterio empleado para separar es una velocidad promedio prevista de 450 km/s, indicando una alta probabilidad de viento rápido a lo largo de una parte significativa del día siguiente, o de picos notables por sí mismos.

2. Información sobre columnas

2.1. Carrington vs 1d

Los atributos de entrada están separados en dos grandes grupos: lecturas a nivel de rotación de Carrington y lecturas de los últimos días.

Las columnas que incluyen la palabra Carrington refieren específicamente a valores desfasados una rotación para capturar el impacto de estructuras persistentes.

Las columnas que incluyen la construcción "1d" refieren al pasado reciente.

El tiempo de la rotación de Carrington se ha fijado a 27,26 días exactos.

2.2. Sufijo $-n$

Todas las columnas referentes a una entrada incluyen un sufijo con formato $-n$. Este sufijo indica el número de saltos de tiempo (días, en este caso) que la entrada correspondiente está separada de la fecha de predicción actual. Por ejemplo, para $n = 2$, la entrada refiere a datos en $t - 2$.

2.3. Unidades y correspondencias

Todos los valores están promediados a nivel diario.

- **Bt**: contiene $|B|$, el módulo de vector de campo magnético medido por MAG. Expresado en nT
- **Bx**: contiene $B_x \in B$. Expresado en nT
- **By**: contiene $B_y \in B$. El sistema de referencia es GSM. Expresado en nT
- **Bz**: contiene $B_z \in B$. El sistema de referencia es GSM. Expresado en nT
- **speed**: contiene la componente radial del vector de velocidad V del viento solar. La dirección siempre es alejándose del Sol. Expresado en km/s
- **p_density**: contiene la densidad de partículas en el viento solar ρ . Expresado en cm^{-3}
- **temperature**: contiene la temperatura del viento solar T . Expresado en K