

Predicción de Tormentas Geomagnéticas con Técnicas de Soporte Vectorial en Regresión

Raquel García Marañón

Tutor: Francisco Javier Yáñez Gestoso

18 de septiembre de 2024



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID



UNIVERSIDAD
COMPLUTENSE
MADRID

Contexto y Motivación [1/2]

- Las **tormentas geomagnéticas** son fenómenos que ocurren cuando el viento solar impacta con la magnetosfera y la ionosfera de la Tierra, provocando una intensificación de la corriente en anillo. Esto se refleja en un aumento notable de la actividad geomagnética. [1]



- Impacto en infraestructuras: afecta a redes eléctricas, sistemas de comunicación y satélites, entre otras.
- Necesidad de modelos predictivos para anticipar eventos y proteger infraestructuras críticas.

Imagen: NOAA SciJinks

Contexto y motivación [2/2]

Definición del Índice Dst

El **índice Dst** mide la intensidad de las perturbaciones geomagnéticas causadas por corrientes anulares durante tormentas geomagnéticas. Se calcula como la media del componente horizontal del campo magnético en cuatro estaciones cercanas al ecuador geomagnético [2]. Expresado en nanoteslas (nT).

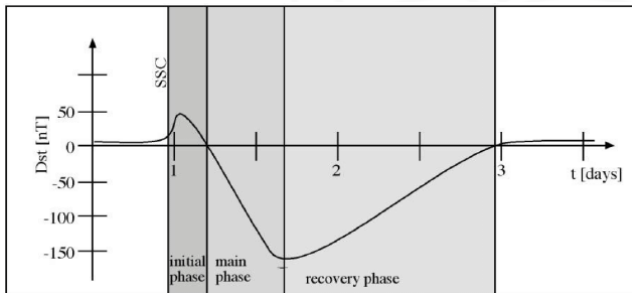


Imagen: Geomagnetic storm phases [3]

Objetivo del Estudio

Objetivo:

- Desarrollar un modelo predictivo para tormentas geomagnéticas utilizando el índice Dst.
 - **Modelo SVR:** Máquina de Vector Soporte en Regresión
- Comparar SVR con modelo alternativo de redes neuronales (**LSTM + MLP**).
- Evaluar la capacidad de predicción a 2, 4 y 6 horas.

Máquinas de Soporte Vectorial para Regresión (SVR)

Modelo SVR

- Aprendizaje automático
- Supervisado
- Regresión
- Encontrar una función que aproxime los datos dentro de un margen de error
 - Penalizar las desviaciones que exceden un umbral equilibrando entre precisión y complejidad
 - Maximizar del margen y minimizar de errores.

Aplicación de la SVR

- Modelo predictivo basado en la definición de un hiperplano en un espacio de alta dimensionalidad.
- Optimización de hiperparámetros clave: C , ϵ y γ .
- Robusto frente a puntos atípicos y adecuado para predicciones robustas en presencia de datos ruidosos.
- Ventaja clave: interpretabilidad frente a modelos de caja negra como las redes neuronales.

Metodología

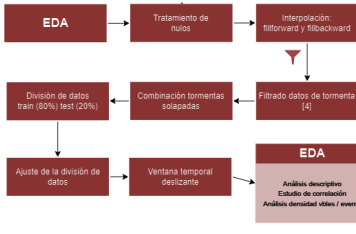
Selección y descarga de datos

Descarga de datos:
ACE y DSCVR
OMNI2_H0_MRG1HR1

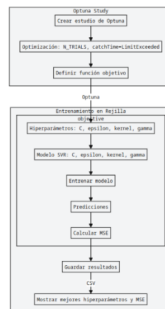
Selección de variables OMNI2

Dst, Bmag, Bc, By_gsw, Bz_gsw, By_gsm, Bz_gsm, P_density,
AP, E_field, plasma_T, plasma_V, Dstime, ID_IMF,
ID_plasma, dev_Bmag, dev_Bc, dev_By, dev_Bz,
dev_P_density, dev_AP, dev_plasma_T

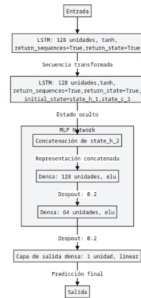
Procesamiento



Entrenamiento



SVR



LSTM-MLP

Metodología: Selección y descarga de datos

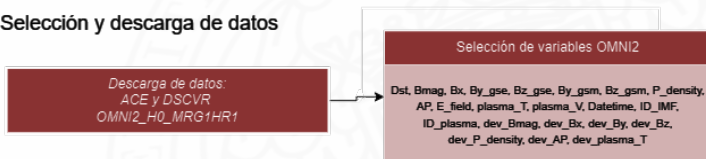
- **Datos** del viento solar y del índice Dst, extraídos de las sondas ACE y DSCOVR

Fuente: OMNI2_H0_MRG1HR1.

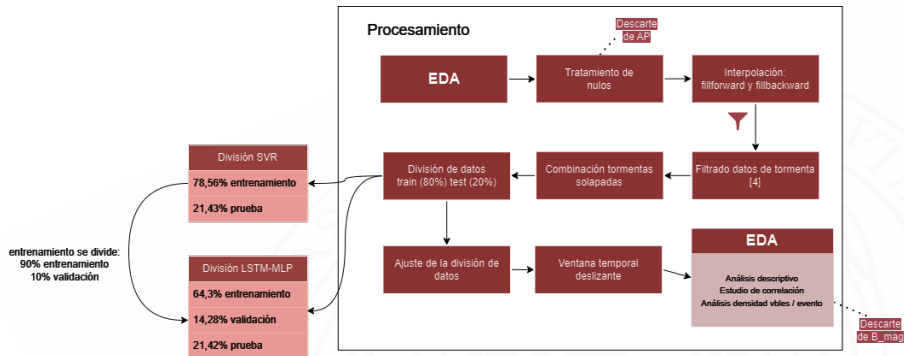


- **Periodo:** 14 de enero de 2001 al 31 de diciembre de 2016 (obtenido de [4]).
- **Variables:**
 - Principales (14): *Dst*, *Bmag*, *Bx*, *By_gse*, *Bz_gse*, *By_gsm*, *Bz_gsm*, *P_density*, *AP*, *E_field*, *plasma_T*, *plasma_V*, *Datetime*
 - Auxiliares: *ID_IMF*, *ID_plasma*, *dev_Bmag*, *dev_Bx*, *dev_By*, *dev_Bz*, *dev_P_density*, *dev_AP*, *dev_plasma_T*

Selección y descarga de datos

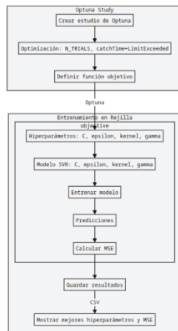


Metodología: Procesamiento de datos

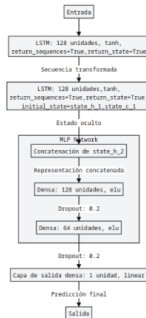


Metodología: Entrenamiento del modelo: ajuste

Entrenamiento



SVR



LSTM-MLP

Resultados y comparaciones

Selección de kernel

kernel	MSE	Nº iteraciones
Linear	3.5076	36
RBF	4.2784	47
Poly	45.7375	51

(a) 2h

kernel	MSE	Nº iteraciones
Linear	5.7601	22
RBF	5.5090	36
Poly	21.4601	39

(b) 4h

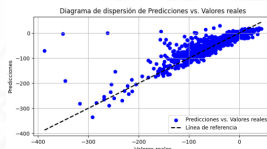
kernel	MSE	Nº iteraciones
Linear	5.3953	27
RBF	6.9295	28
Poly	14.6611	37

(c) 6h

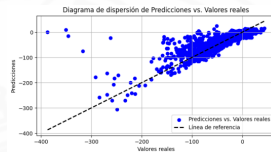
Resultados Kernel lineal



(a) 2h



(b) 4h



(c) 6h

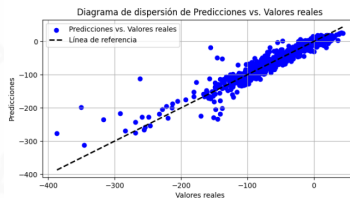
Métrica	RMSE	MSE	MAE	R ²	MedAE	Var. expl.	Max error
Valor (2h)	11.11	123.35	6.11	0.9225	3.87	0.92	193.33
Valor (4h)	18.85	355.24	9.41	0.7778	5.66	0.78	348.96
Valor (6h)	23.55	554.80	11.68	0.6541	6.60	0.66	387.14

Resultados y Comparaciones a 2h [1/3]

Métrica	RMSE	MSE	MAE	R^2	MedAE	Varianza explicada	Max error
SVR (2h)	11.11	123.35	6.11	0.9225	3.87	0.92	193.33
LSTM-MLP (2h)	11.32	128.20	6.22	0.9195	3.91	0.9196	152.47



(a) SVR

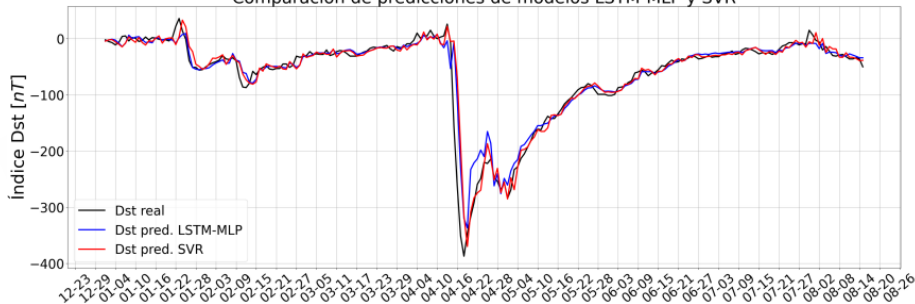


(b) LSTM-MLP

Resultados y Comparaciones a 2h [2/3]

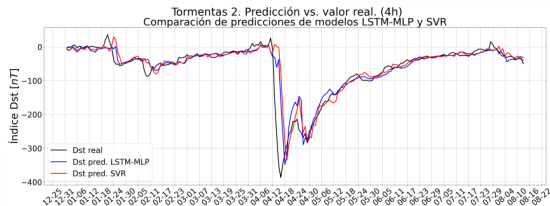
	RMSE	MSE	MAE	R ²	MedAE	Var. expl.	Max error
LSTM	21.635	468.073	10.206	0.922	5.410	0.924	153.092
SVR	22.348	499.436	9.546	0.917	5.106	0.917	193.334

Tormentas 2. Predicción vs. valor real. (2h)
Comparación de predicciones de modelos LSTM-MLP y SVR

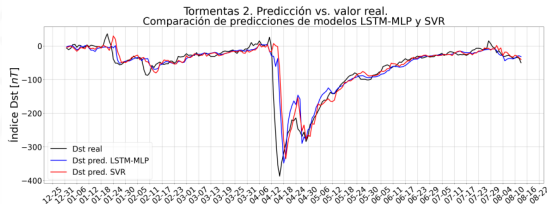


Resultados y Comparaciones a 4h y 6h [3/3]

Métrica	LSTM	SVR
RMSE	21.635	22.348
MSE	468.073	499.436
MAE	10.206	9.546
R ²	0.922	0.917
MedAE	5.410	5.106
Var. expl.	0.924	0.917
Max error	153.092	193.334



Métrica	LSTM	SVR
RMSE	21.635	22.348
MSE	468.073	499.436
MAE	10.206	9.546
R ²	0.922	0.917
MedAE	5.410	5.106
Var. expl.	0.924	0.917
Max error	153.092	193.334



Conclusiones

- SVR es eficaz en la predicción de tormentas geomagnéticas, con resultados precisos y consistentes.
- Aprendizajes para trabajos posteriores: Entrenamiento demasiado intensivo. Necesidad de gestionar adecuadamente los recursos disponibles y el alcance del proyecto.
- Ventaja de SVR sobre LSTM-MLP en interpretabilidad y robustez frente a datos ruidosos.
- No se observa una mejora significativa sobre algunos modelos de la bibliografía.
- El enfoque basado en SVR sigue siendo prometedor para objetivos científicos, pero no necesariamente el más eficiente para todas las aplicaciones. El SVR podría quedarse rezagado frente a métodos más avanzados, como redes neuronales profundas, que son capaces de manejar grandes volúmenes de datos en tiempo real.

Referencias I

- [1] W. Gonzalez, J. Joselyn, Y. Kamide, H. Kroehl, G. Rostoker, B. Tsurutani, and V. Vasyliunas. What is a geomagnetic storm. *Journal of Geophysical Research*, 99:5771–5792, 1994. doi: 10.1029/93JA02867.
- [2] F. Nuraeni, M. Ruhimat, M. A. Aris, E. A. Ratnasari, and C. Purnomo. Development of 24 hours dst index prediction from solar wind data and imf bz using narx. *Journal of Physics: Conference Series*, 2214, 2022. doi: 10.1088/1742-6596/2214/1/012024.
- [3] Farah Mohd Kasran, Jusoh Huzaimy, Siti Amalina Enche Ab Rahim, and Noradlina Abdullah. Geomagnetically induced currents (gics) in equatorial region. pages 112–117, 10 2018. doi: 10.1109/ICSEngT.2018.8606391.
- [4] A. Gruet, E. Camporeale, and B. Swiger. Multi-hour-ahead dst index prediction using multi-fidelity boosted neural networks. *Journal of Geophysical Research: Space Physics*, 2018. doi: 10.48550/arXiv.2209.12571.