

Análisis de Resultados: Regresión KNN vs. SVM para Pronóstico de Series Temporales

1. Métricas de Desempeño

K-Nearest Neighbors (KNN):

Variable	R ²	RMSE	MAE
generation.solar	0.292	1556.813	1101.061
generation.wind.onshore	-0.906	3953.569	3243.733
total.load.actual	-0.228	4468.826	3584.862

Support Vector Machines (SVM):

Variable	R ²	RMSE	MAE
generation.solar	0.503	1304.703	966.055
generation.wind.onshore	-0.698	3731.805	2952.332
total.load.actual	-0.334	4657.628	3913.831

2. Interpretación de Métricas

- **R² (Coeficiente de Determinación):**
 - **Valores positivos (ej. solar con SVM: 0.503):** El modelo explica parcialmente la varianza de los datos. SVM supera a KNN en generación solar.
 - **Valores negativos (ej. eólica en KNN: -0.906):** Indica que el modelo es peor que predecir la media constante. Esto sugiere un **sobreajuste severo** o **inadecuación del modelo** para capturar patrones en estas variables.
- **RMSE y MAE:**

- **Generación solar:** SVM reduce el error (RMSE: 1304 vs. 1556 en KNN), lo que confirma su mejor desempeño.
- **Eólica y carga total:** Ambos modelos tienen errores altos, pero SVM es ligeramente mejor en eólica (RMSE: 3731 vs. 3953 en KNN).

3. Análisis de Gráficos

KNN:

- **Generación solar:** Muestra cierta correlación con la tendencia real, pero con alta varianza (picos erráticos).
- **Eólica terrestre:** Pronósticos muy dispersos, sin seguir la tendencia real (coherente con R^2 negativo).
- **Carga total:** Aunque el R^2 es menos negativo, los errores absolutos (MAE \approx 3585) indican predicciones poco confiables.

SVM:

- **Generación solar:** Línea más suave y cercana a la tendencia real, reflejando el mayor R^2 .
- **Eólica y carga total:** Aunque los errores son menores que en KNN, las predicciones aún se desvían significativamente de los valores reales.

4. Causas de los Resultados

- **Generación solar:**
 - **Patrones más predecibles:** La generación solar tiene ciclos diarios y estacionales claros (día/noche, verano/invierno), lo que favorece a modelos como SVM con kernel no lineal.
 - **KNN:** Sensible a ruido en datos históricos, lo que explica su mayor varianza.
- **Eólica terrestre y carga total:**
 - **Alta variabilidad:** La generación eólica depende de factores meteorológicos complejos (vientos erráticos), y la carga total puede verse afectada por eventos no estacionales (ej. festivos).

- **Falta de variables explicativas:** Ambos modelos usan solo datos históricos de la misma variable. Incorporar variables externas (ej. temperatura, velocidad del viento) podría mejorar los resultados.

5. Recomendaciones

1. Mejorar la Ingeniería de Características:

- Añadir variables externas (meteorológicas, festivos) para capturar mejor los patrones.
- Probar ventanas de tiempo más largas (ej. 30 días) o transformaciones (diferenciación, normalización).

2. Ajustar Hiperparámetros:

- **KNN:** Optimizar el número de vecinos (k) mediante validación cruzada.
- **SVM:** Probar otros kernels (ej. polinomial) o ajustar parámetros de regularización (C, gamma).

3. Explorar Otros Modelos:

- Modelos basados en árboles (Random Forest, XGBoost) para manejar no linealidades.
- Redes Neuronales Recurrentes (LSTM) para dependencias temporales complejas.

4. Validar con Más Datos:

- Asegurar que el conjunto de entrenamiento cubra múltiples estacionalidades (ej. varios años).

6. Conclusión

- **SVM es superior a KNN** para generación solar, gracias a su capacidad para modelar relaciones no lineales.
- Ambos modelos **fracasan en eólica y carga total**, probablemente por la naturaleza caótica de estas variables y la falta de predictores relevantes.
- Se requieren mejoras en el preprocesamiento y la inclusión de variables externas para obtener pronósticos confiables en todas las variables.