

Marco Teórico: Support Vector Regression (SVR)

El Support Vector Regression (SVR) es una técnica de aprendizaje supervisado que extiende el algoritmo de Support Vector Machines (SVM) al dominio de la regresión. El objetivo principal del SVR es encontrar una función que ajuste los datos dentro de un margen definido por un parámetro de tolerancia (epsilon), minimizando el error y la complejidad del modelo.

Kernels en SVR

El SVR utiliza funciones kernel para transformar los datos originales a un espacio de mayor dimensionalidad donde es más fácil realizar ajustes. Los kernels más comunes son:

- Kernel Lineal: Captura relaciones lineales.
- Kernel Polinomial: Captura relaciones no lineales de grado especificado.
- Kernel RBF (Radial Basis Function): Es altamente efectivo para datos complejos y no lineales.

El kernel adecuado depende de la naturaleza de los datos y la relación entre las variables.

Parámetros importantes en SVR

- C: Controla la penalización por errores. Valores altos de C ajustan más los datos, pero pueden llevar al sobreajuste. Valores bajos generan modelos más simples.
- Epsilon: Define el margen de tolerancia para errores. Los errores dentro de este margen no son penalizados, lo que permite un ajuste más robusto.
- Gamma (solo para kernels no lineales): Controla la influencia de un solo dato en el modelo. Valores altos de gamma generan curvas ajustadas a los datos cercanos; valores bajos generan curvas más suaves.

Vectores de soporte

Los vectores de soporte son los puntos de datos más cercanos al margen de tolerancia ($f(x) \pm \epsilon$). Estos puntos son cruciales porque determinan la posición del hiperplano de ajuste y el margen. Los vectores de soporte son fundamentales tanto para la clasificación como para la

regresión.

Comparación entre Kernels

- Kernel Lineal: Ideal para datos con relaciones lineales simples. Rápido y eficiente.
- Kernel Polinomial: Útil para datos no lineales con relaciones polinómicas de grado conocido.
- Kernel RBF: Más versátil y comúnmente utilizado. Puede capturar relaciones complejas y no lineales.

Aplicaciones de SVR

- Predicción de series temporales.
- Modelado de relaciones no lineales en datos científicos.
- Ajuste de curvas en problemas de ingeniería.
- Predicción de precios en mercados financieros.

Conclusión

El Support Vector Regression es una herramienta poderosa para resolver problemas de regresión. La selección de parámetros como C , epsilon y el kernel adecuado es crucial para obtener un modelo eficiente. Con un buen entendimiento de los datos, el SVR puede ser aplicado exitosamente en una amplia variedad de problemas prácticos.